

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2005, 09:24:38 ; Search time 70 Seconds  
(without alignments)  
1055.303 Million cell updates/sec

Title: US-09-084-691B-206  
Perfect score: 1045  
Sequence: 1 MSTLPKPKQKTKRNTNRRPT.....CSFSIFLLALLSCLTTPASA 191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1045	100.0	191	2	AAR92987 Hepatitis
2	996	95.3	191	2	AAR92972 Hepatitis
3	982	94.0	191	2	AAR92974 Hepatitis
4	982	94.0	191	2	AAR92973 Hepatitis
5	981	93.9	191	2	AAR92953 Hepatitis
6	980	93.8	191	2	AAR92978 Hepatitis
7	978	93.6	191	2	AAR92976 Hepatitis
8	977	93.5	191	2	AAR92977 Hepatitis
9	977	93.5	196	2	AAR74048 Synthetic
10	977	93.5	196	2	AAR74047 Synthetic
11	977	93.5	319	2	AAR96546 Hepatitis
12	977	93.5	326	2	AAR22137 HCV-HC59
13	977	93.5	733	2	AAR38278 NANB hepa
14	977	93.5	2894	2	AAR24440 Composite
15	977	93.5	2894	2	AAR70230 Composite
16	977	93.5	3011	2	AAR66995 Hepatitis
17	976	93.4	191	2	AAR44010 Hepatitis
18	976	93.4	191	2	AAR92938 Hepatitis
19	976	93.4	191	2	AAR92941 Hepatitis
20	976	93.4	191	2	AAR92937 Hepatitis
21	976	93.4	191	2	AAR92939 Hepatitis
22	976	93.4	191	2	AAR92940 Hepatitis
23	976	93.4	191	3	AAY94411 Human hep
24	976	93.4	249	8	ADJ10436 HCV Core
25	976	93.4	473	8	ADJ10438 HCV Core (

26	976	93.4	502	2	AAR67591 Hepatitis
27	976	93.4	967	2	AAR79222 pHCV141-e
28	976	93.4	1006	2	AAW12715 HCV genom
29	976	93.4	1648	4	AAR79221 pHCV176-e
30	976	93.4	2984	4	AAE00449 Hepatitis
31	976	93.4	2984	4	AAE00447 Hepatitis
32	976	93.4	2984	4	AAE00442 Hepatitis
33	976	93.4	3002	7	ADM24822 Hepatitis
34	976	93.4	3011	2	AAR40119 HCV genom
35	976	93.4	3011	2	AAR40120 HCV genom
36	976	93.4	3011	2	AAR79232 HCV seque
37	976	93.4	3011	2	AAW77397 Hepatitis
38	976	93.4	3011	2	AAW77398 Hepatitis
39	976	93.4	3011	2	AAW98020 Infectiou
40	976	93.4	3011	4	AAE59173 Protein e
41	976	93.4	3011	4	AAE31169 Amino aci
42	976	93.4	3011	5	AAU84597 HCV polyp
43	976	93.4	3011	5	AAU79221 Hepatitis
44	976	93.4	3011	5	AAE19888 Hepatitis
45	976	93.4	3011	6	ABP71460 Amino aci

## ALIGNMENTS

RESULT 1  
AAR92987  
ID AAR92987 standard; protein; 191 AA.  
XX  
AC AAR92987;  
XX  
DT 02-OCT-1996 (first entry)  
XX  
DE Hepatitis C virus isolate HK2 core protein.  
XX  
KW HCV; E; envelope 1; core protein; HCV genotyping; antibody; vaccine;  
KW Hepatitis.  
XX  
OS Hepatitis C virus.  
XX  
PN WO9605315-A2.  
XX  
PD 22-FEB-1996.  
XX  
PF 15-AUG-1995; 95WO-US010398.  
XX  
PR 15-AUG-1994; 94US-00290665.  
XX  
(USSH ) US SEC DEPT HEALTH.  
XX  
Bukh J, Miller RH, Purcell RH;  
DR WPI; 1996-139709/14.  
DR N-PSDB; AAR16661.  
XX  
DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection.  
XX  
Claim 4; Page 223; 340pp; English.  
XX  
AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection  
XX  
SQ Sequence 191 AA;

Query Match 100.0%; Score 1045; DB 2; Length 191;  
Best Local Similarity 100.0%; Pred. No. 2.6e-90;

Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTLPKPKQKRTNRRPTDVKPPGGQIVGGVYLLPRGPRGLGVRATRTKTSERSQPRG 60  
 Db 1 MSTLPKPKQKRTNRRPTDVKPPGGQIVGGVYLLPRGPRGLGVRATRTKTSERSQPRG 60

Qy 61 RROPIPKARQPGRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGPNDPRRSRNLG 120  
 Db 61 RROPIPKARQPGRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGPNDPRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPVWGAPLGVAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180  
 Db 121 KVIDTLTCGFADLMGYIPVWGAPLGVAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180

Qy 181 LLCLTTPASA 191  
 Db 181 LLCLTTPASA 191

RESULT 2  
 AAR92972  
 ID AAR92972 standard; protein; 191 AA.  
 AC AAR92972;  
 XX  
 DT 02-OCT-1996 (first entry)  
 DE Hepatitis C virus isolate Z4 core protein.  
 KW HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;  
 KW hepatitis.  
 XX  
 OS Hepatitis C virus.  
 PN WO9605315-A2.  
 PD 22-FEB-1996.  
 XX  
 PF 15-AUG-1995; 95WO-US010398.  
 XX  
 PR 15-AUG-1994; 94US-00290665.  
 XX  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Bukh J, Miller RH, Purcell RH;  
 XX WPI; 1996-139709/14.  
 DR N-PSDB; AAT16646.  
 XX  
 PT DNA and amino acid sequence of HCV envelope 1 and core proteins - used to  
 PT determine HCV genotype and as vaccines against HCV infection.  
 XX  
 PS Claim 4; Page 210-211; 340pp; English.  
 XX  
 CC AAR92936-R92987 are HCV core proteins derived from 52 different HCV  
 CC isolates. Isolated cDNA sequences are used for the prodn. of primers  
 CC useful for detecting the presence of HCV in a sample, the primers are  
 CC also useful for HCV genotyping. Proteins encoded by the cDNAs can be used  
 CC in vaccines for immunising against HCV infection. The proteins may also  
 CC be used to detect antibodies against HCV in serum, saliva, lymphocytes or  
 CC other mononuclear cells. The antibodies may be used in the prevention of  
 CC HCV infection  
 XX  
 SQ Sequence 191 AA;  
 Query Match 95.3%; Score 996; DB 2; Length 191;  
 Best Local Similarity 94.8%; Pred. No. 1.1e-85;  
 Matches 181; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MSTLPKPKQKRTNRRPTDVKPPGGQIVGGVYLLPRGPRGLGVRATRTKTSERSQPRG 60  
 Db 1 MSTLPKPKQKRTNRRPTDVKPPGGQIVGGVYLLPRGPRGLGVRATRTKTSERSQPRG 60

Qy 61 RROPIPKARQPGRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGPNDPRRSRNLG 120  
 Db 61 RROPIPKARQPGRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGPNDPRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPVWGAPLGVAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180  
 Db 121 KVIDTLTCGFADLMGYIPVWGAPLGVAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180

Qy 181 LLCLTTPASA 191  
 Db 181 LLCLTTPASA 191

RESULT 3  
 AAR92974  
 ID AAR92974 standard; protein; 191 AA.  
 AC AAR92974;  
 XX  
 DT 02-OCT-1996 (first entry)  
 DE Hepatitis C virus isolate Z1 core protein.  
 KW HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;  
 KW hepatitis.  
 XX  
 OS Hepatitis C virus.  
 PN WO9605315-A2.  
 PD 22-FEB-1996.  
 XX  
 PF 15-AUG-1995; 95WO-US010398.  
 XX  
 PR 15-AUG-1994; 94US-00290665.  
 XX  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Bukh J, Miller RH, Purcell RH;  
 XX WPI; 1996-139709/14.  
 DR N-PSDB; AAT16648.  
 XX  
 PT DNA and amino acid sequence of HCV envelope 1 and core proteins - used to  
 PT determine HCV genotype and as vaccines against HCV infection.  
 XX  
 PS Claim 4; Page 212; 340pp; English.  
 XX  
 CC AAR92936-R92987 are HCV core proteins derived from 52 different HCV  
 CC isolates. Isolated cDNA sequences are used for the prodn. of primers  
 CC useful for detecting the presence of HCV in a sample, the primers are  
 CC also useful for HCV genotyping. Proteins encoded by the cDNAs can be used  
 CC in vaccines for immunising against HCV infection. The proteins may also  
 CC be used to detect antibodies against HCV in serum, saliva, lymphocytes or  
 CC other mononuclear cells. The antibodies may be used in the prevention of  
 CC HCV infection  
 XX  
 SQ Sequence 191 AA;  
 Query Match 94.0%; Score 982; DB 2; Length 191;  
 Best Local Similarity 93.7%; Pred. No. 2.2e-84;  
 Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MSTLPKPKQKRTNRRPTDVKPPGGQIVGGVYLLPRGPRGLGVRATRTKTSERSQPRG 60  
 Db 1 MSTLPKPKQKRTNRRPTDVKPPGGQIVGGVYLLPRGPRGLGVRATRTKTSERSQPRG 60

Qy 61 RROPIPKARQPGRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGPNDPRRSRNLG 120  
 Db 61 RROPIPKARQPGRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGPNDPRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPVWGAPLGVAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180  
 Db 121 KVIDTLTCGFADLMGYIPVWGAPLGVAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180

```

Db 121 KVIDLTCTGCFADLMGYIPLVGAPVGGVARALAHGVRAVEDGINYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
  |||||
Db 181 LLSCLTTPASA 191

RESULT 4
AAR92973
ID AAR92973 standard; protein; 191 AA.
XX
AC AAR92973;
XX
DT 02-OCT-1996 (first entry)
DE Hepatitis C virus isolate 28 core protein.
XX
KW HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
KW hepatitis.
XX
OS Hepatitis C virus.
XX
PN WO9605315-A2.
XX
PD 22-FEB-1996.
XX
PF 15-AUG-1995; 95WO-US010398.
XX
PR 15-AUG-1994; 94US-00290665.
XX
PA (USSH ) US SEC DEPT HEALTH.
XX
PI Bukh J, Miller RH, Purcell RH;
XX
WPI: 1996-139709/14.
DR N-PSDB; AAT16647.
XX
DNA and amino acid sequence of HCV envelope 1 and core proteins - used to
determine HCV genotype and as vaccines against HCV infection.
XX
Claim 4; Page 211-212; 340pp; English.
XX
AAR92936-R92987 are HCV core proteins derived from 52 different HCV
isolates. Isolated cDNA sequences are used for the prodn. of primers
useful for detecting the presence of HCV in a sample, the primers are
also useful for HCV genotyping. Proteins encoded by the cDNAs can be used
in vaccines for immunising against HCV infection. The proteins may also
be used to detect antibodies against HCV in serum, saliva, lymphocytes or
other mononuclear cells. The antibodies may be used in the prevention of
HCV infection
XX
SQ Sequence 191 AA;
Query Match 94.0%; Score 982; DB 2; Length 191;
Best Local Similarity 93.7%; Pred. No. 2.2e-84;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 1 MSTLPKPKQKRNTRRPTDVKFPGGQIVGGVLLPRRGVRLGVRATKTSERSQPRG 60
  |||||
Db 1 MSTNPKPQKRNTRRPMVDFKPGGQIVGGVLLPRRGVRLGVRATKTSERSQPRG 60
  |||||
QY 61 RRQPIPKARQPGQRHWAQGYWPVLYGNEGCGWAGWLLSPRGRPHWGPNDRPRRSRLNG 120
  |||||
Db 61 RRQPIPKARREGRSQAQGYWPVLYGNEGCGWAGWLLSPRGRPSWGNDRPRRSRLNG 120
  |||||
QY 121 KVIDLTCTGCFADLMGYIPLVGAPVGGVARALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
  |||||
Db 121 KVIDLTCTGCFADLMGYIPLVGAPVGGVARALAHGVRAVEDGINYATGNLPGCSFSIFLLA 180
  |||||
QY 181 LLSCLTTPASA 191
  |||||
Db 181 LLSCLTTPASA 191
  |||||

```

```

RESULT 5
AAR92953
ID AAR92953 standard; protein; 191 AA.
XX
AC AAR92953;
XX
DT 02-OCT-1996 (first entry)
DE Hepatitis C virus isolate HK3 core protein.
XX
KW HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
KW hepatitis.
XX
OS Hepatitis C virus.
XX
PN WO9605315-A2.
XX
PD 22-FEB-1996.
XX
PF 15-AUG-1995; 95WO-US010398.
XX
PR 15-AUG-1994; 94US-00290665.
XX
PA (USSH ) US SEC DEPT HEALTH.
XX
PI Bukh J, Miller RH, Purcell RH;
XX
WPI: 1996-139709/14.
DR N-PSDB; AAT16627.
XX
DNA and amino acid sequence of HCV envelope 1 and core proteins - used to
determine HCV genotype and as vaccines against HCV infection.
XX
Claim 4; Page 194-195; 340pp; English.
XX
AAR92936-R92987 are HCV core proteins derived from 52 different HCV
isolates. Isolated cDNA sequences are used for the prodn. of primers
useful for detecting the presence of HCV in a sample, the primers are
also useful for HCV genotyping. Proteins encoded by the cDNAs can be used
in vaccines for immunising against HCV infection. The proteins may also
be used to detect antibodies against HCV in serum, saliva, lymphocytes or
other mononuclear cells. The antibodies may be used in the prevention of
HCV infection
XX
SQ Sequence 191 AA;
Query Match 93.9%; Score 981; DB 2; Length 191;
Best Local Similarity 93.7%; Pred. No. 2.8e-84;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 1 MSTLPKPKQKRNTRRPTDVKFPGGQIVGGVLLPRRGVRLGVRATKTSERSQPRG 60
  |||||
Db 1 MSTNPKPQKRNTRRPMVDFKPGGQIVGGVLLPRRGVRLGVRATKTSERSQPRG 60
  |||||
QY 61 RRQPIPKARQPGQRHWAQGYWPVLYGNEGCGWAGWLLSPRGRPHWGPNDRPRRSRLNG 120
  |||||
Db 61 RRQPIPKARQPEGRWAQGYWPVLYGNEGCGWAGWLLSPRGRPNWGTDPDRRSRLNG 120
  |||||
QY 121 KVIDLTCTGCFADLMGYIPLVGAPVGGVARALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
  |||||
Db 121 KVIDLTCTGCFADLMGYIPLVGAPVGGVARALAHGVRLVEDGVNATGNLPGCSFSIFLLA 180
  |||||
QY 181 LLSCLTTPASA 191
  |||||
Db 181 LLSCLTTPASA 191
  |||||

RESULT 6
AAR92978
ID AAR92978 standard; protein; 191 AA.
XX
AC AAR92978;

```

XX 02-OCT-1996 (first entry)  
 XX Hepatitis C virus isolate DK13 core protein.  
 DE HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;  
 XX hepatitis.  
 XX Hepatitis C virus.  
 XX WO9605315-A2.  
 XX 22-FEB-1996.  
 XX 15-AUG-1995; 95WO-US010398.  
 XX 15-AUG-1994; 94US-00290665.  
 XX (USSH ) US SEC DEPT HEALTH.  
 XX Bukh J, Miller RH, Purcell RH;  
 PI WPI; 1996-139709/14.  
 DR N-PSDB; AAT16652.  
 XX DNA and amino acid sequence of HCV envelope 1 and core proteins - used to  
 PT determine HCV genotype and as vaccines against HCV infection.  
 XX Claim 4; Page 215-216; 340pp; English.  
 XX AAR92936-R92987 are HCV core proteins derived from 52 different HCV  
 CC isolates. Isolated cDNA sequences are used for the prodn. of primers  
 CC also useful for detecting the presence of HCV in a sample, the primers are  
 CC in vaccines for immunising against HCV infection. The proteins can be used  
 CC be used to detect antibodies against HCV in serum, saliva, lymphocytes or  
 CC other mononuclear cells. The antibodies may be used in the prevention of  
 CC HCV infection  
 XX Sequence 191 AA;  
 SQ Query Match 93.8%; Score 980; DB 2; Length 191;  
 Best Local Similarity 93.7%; Pred. No. 3.5e-84;  
 Matches 179; Conservative 4; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 MSTLPKQKTKRNTNRRPTDVKFPGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60  
 DB 1 MSTNPKQKTKRNTNRRPMDVKFPGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60  
 QY 61 RROPIPKARQPGQRHWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPHWGPNDPRRSRLG 120  
 DB 61 RROPIPKARQPGQRHWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPHWGPNDPRRSRLG 120  
 QY 121 KVIDTLTCGFADLMGYIPVWGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180  
 DB 121 KVIDTLTCGFADLMGYIPVWGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180  
 QY 181 LLSCLTTPASA 191  
 DB 181 LLSCLTTPASA 191  
 RESULT 7  
 AAR92976  
 ID AAR92976 standard; protein; 191 AA.  
 XX AAR92976;  
 XX 02-OCT-1996 (first entry)  
 DE Hepatitis C virus isolate Z6 core protein.  
 XX HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;  
 KW hepatitis.

KW hepatitis.  
 XX Hepatitis C virus.  
 XX WO9605315-A2.  
 XX 22-FEB-1996.  
 XX 15-AUG-1995; 95WO-US010398.  
 XX 15-AUG-1994; 94US-00290665.  
 XX (USSH ) US SEC DEPT HEALTH.  
 XX Bukh J, Miller RH, Purcell RH;  
 PI WPI; 1996-139709/14.  
 DR N-PSDB; AAT16650.  
 XX DNA and amino acid sequence of HCV envelope 1 and core proteins - used to  
 PT determine HCV genotype and as vaccines against HCV infection.  
 XX Claim 4; Page 214; 340pp; English.  
 XX AAR92936-R92987 are HCV core proteins derived from 52 different HCV  
 CC isolates. Isolated cDNA sequences are used for the prodn. of primers  
 CC also useful for detecting the presence of HCV in a sample, the primers are  
 CC in vaccines for immunising against HCV infection. The proteins can be used  
 CC be used to detect antibodies against HCV in serum, saliva, lymphocytes or  
 CC other mononuclear cells. The antibodies may be used in the prevention of  
 CC HCV infection  
 XX Sequence 191 AA;  
 SQ Query Match 93.6%; Score 978; DB 2; Length 191;  
 Best Local Similarity 93.2%; Pred. No. 5.3e-84;  
 Matches 178; Conservative 5; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 MSTLPKQKTKRNTNRRPTDVKFPGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60  
 DB 1 MSTNPKQKTKRNTNRRPMDVKFPGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60  
 QY 61 RROPIPKARQPGQRHWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPHWGPNDPRRSRLG 120  
 DB 61 RROPIPKARQPGQRHWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPHWGPNDPRRSRLG 120  
 QY 121 KVIDTLTCGFADLMGYIPVWGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180  
 DB 121 KVIDTLTCGFADLMGYIPVWGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180  
 QY 181 LLSCLTTPASA 191  
 DB 181 LLSCLTTPASA 191  
 RESULT 8  
 AAR92977  
 ID AAR92977 standard; protein; 191 AA.  
 XX AAR92977;  
 XX 02-OCT-1996 (first entry)  
 DE Hepatitis C virus isolate Z7 core protein.  
 XX HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;  
 KW hepatitis.  
 XX Hepatitis C virus.  
 XX WO9605315-A2.

PD 22-FEB-1996.  
XX PF 15-AUG-1995; 95WO-US010398.  
XX PR 15-AUG-1994; 94US-00290665.  
XX (USSH ) US SEC DEPT HEALTH.  
XX PA Bukh J, Miller RH, Purcell RH;  
XX PI WPI; 1996-139709/14.  
XX DR N-PSDB; AAT16651.  
XX PT DNA and amino acid sequence of HCV envelope 1 and core proteins - used to  
XX determine HCV genotype and as vaccines against HCV infection.  
XX PS Claim 4; Page 214-215; 340pp; English.  
XX CC AAR92936-R92987 are HCV core proteins derived from 52 different HCV  
XX isolates. Isolated cDNA sequences are used for the prodn. of primers  
XX useful for detecting the presence of HCV in a sample, the primers are  
XX also useful for HCV genotyping. Proteins encoded by the cDNAs can be used  
XX in vaccines for immunising against HCV infection. The proteins may also  
XX be used to detect antibodies against HCV in serum, saliva, lymphocytes or  
XX other mononuclear cells. The antibodies may be used in the prevention of  
XX HCV infection  
XX  
XX Sequence 191 AA;  
Query Match 93.5%; Score 977; DB 2; Length 191;  
Best Local Similarity 93.2%; Pred. No. 6.6e-84;  
Matches 178; Conservative 5; Mismatches 8; Indels 0; Gaps 0;  
QY 1 MSTLPKPKQKTRNTNRRPTDVKFPGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60  
DB 1 MSTNPKPKQKTRNTNRRPMDVKFPGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60  
QY 61 RRQIPKARQPGGRHWAQPGYPWPPLYGNEGCGWAGWLLSPRGRSPHWPNDPRRSRNLG 120  
DB 61 RRQIPKARRSGRSGWAQPGYPWPPLYGNEGCGWAGWLLSPRGRSPHWPNDPRRSRNLG 120  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRALEDGINYATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPVVGAPVGGVARALAHGVRALEDGINYATGNLPGCSFSIFLLA 180  
QY 181 LLSCLTTPASA 191  
DB 181 LLSCLTVPASA 191  
RESULT 9  
AAR74048  
ID AAR74048 standard; protein; 196 AA.  
XX AC AAR74048;  
XX 25-MAR-2003 (revised)  
DT 26-NOV-1995 (first entry)  
XX XX  
DE Synthetic HCV nucleocapsid protein.  
XX Hepatitis C virus; vector; expression; prokaryotic cells; detection;  
KW diagnosis.  
XX Synthetic.  
XX WO9511980-A2.  
XX 04-MAY-1995.  
XX 25-OCT-1994; 94WO-US012166.  
XX 25-OCT-1993; 93US-00141917.  
XX (USSH ) US SEC DEPT HEALTH & HUMAN SERVICES.  
XX PI Khudyakov Y, Fields HA;  
XX DR WPI; 1995-178872/23.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX PA Khudyakov Y, Fields HA;  
XX PI WPI; 1995-178872/23.  
XX DR N-PSDB; AAR92071.  
XX PT Vector for expression of a synthetically produced protein coding DNA -  
XX pref. encodes the hepatitis C virus nucleocapsid protein which can be  
XX used in the detection of HCV antibodies.  
XX PS Disclosure; Page 37; 44pp; English.  
XX CC The sequence is that of a synthetic nucleotide sequence encoding the  
XX hepatitis C virus nucleocapsid protein. The gene is positioned in a  
XX vector for efficient expression in prokaryotic cells. The expressed  
XX protein can be used in tests for the detection of antibodies specific for  
XX the HCV C protein. See also AAR74047. (Updated on 25-MAR-2003 to correct  
XX PN field.)  
XX Sequence 196 AA;  
Query Match 93.5%; Score 977; DB 2; Length 196;  
Best Local Similarity 92.1%; Pred. No. 6.8e-84;  
Matches 176; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
QY 1 MSTLPKPKQKTRNTNRRPTDVKFPGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60  
DB 1 MSTIPKPKQKTRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60  
QY 61 RRQIPKARQPGGRHWAQPGYPWPPLYGNEGCGWAGWLLSPRGRSPHWPNDPRRSRNLG 120  
DB 61 RRQIPKVRREGRWTAQPGYPWPPLYGNEGCGWAGWLLSPRGRSPHWPNDPRRSRNLG 120  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRALEDGINYATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGAARALAHGVRALEDGINYATGNLPGCSFSIFLLA 180  
QY 181 LLSCLTTPASA 191  
DB 181 LLSCLTVPASA 191  
RESULT 10  
AAR74047  
ID AAR74047 standard; protein; 196 AA.  
XX AC AAR74047;  
XX 25-MAR-2003 (revised)  
DT 26-NOV-1995 (first entry)  
XX XX  
DE Synthetic HCV nucleocapsid protein.  
XX Hepatitis C virus; vector; expression; prokaryotic cells; detection;  
KW diagnosis.  
XX Synthetic.  
XX WO9511980-A2.  
XX 04-MAY-1995.  
XX 25-OCT-1994; 94WO-US012166.  
XX 25-OCT-1993; 93US-00141917.  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX PI Khudyakov Y, Fields HA;  
XX DR WPI; 1995-178872/23.

DR N-PSDB; AAQ92070.  
 XX Vector for expression of a synthetically produced protein coding DNA -  
 PT pref. encodes the hepatitis C virus nucleocapsid protein which can be  
 PT used in the detection of HCV antibodies.  
 XX  
 XX Disclosure; Page 35; 44pp; English.  
 PS  
 CC The sequence is that of a synthetic nucleotide sequence encoding the  
 CC hepatitis C virus nucleocapsid protein. The gene is positioned in a  
 CC vector for efficient expression in prokaryotic cells. The expressed  
 CC protein can be used in tests for the detection of antibodies specific for  
 CC the HCV C protein. See also AAR74048. (Updated on 25-MAR-2003 to correct  
 CC PN field.)  
 XX  
 XX Sequence 196 AA;  
 SQ  
 Query Match 93.5%; Score 977; DB 2; Length 196;  
 Best Local Similarity 92.1%; Pred. No. 6.8e-84;  
 Matches 176; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 MSTLPKQKTKRNTNRPTDVKFPGGGQIVGGVYLLPRGPRLGVRATRTKTSERSQPRG 60  
 DB 1 MSTLPKQKTKRNTNRPTDVKFPGGGQIVGGVYLLPRGPRLGVRATRTKTSERSQPRG 60  
 QY 61 RROPIKARQPOGRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGPNDRPRRSNGLG 120  
 DB 61 RROPIKVRPRGRTWAQPGYPWPLYGNEGCGWAGWLLSPRGRPSWGTDPDRRSNGLG 120  
 QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180  
 DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGVNATGNLPGCSFSIFLLA 180  
 QY 181 LLSCLTTPASA 191  
 DB 181 LLSCLTTPASA 191  
 RESULT 11  
 AAR96546  
 ID AAR96546 standard; peptide; 319 AA.  
 AC AAR96546;  
 XX  
 DT 10-MAR-1997 (first entry)  
 XX  
 DE Hepatitis C virus types 9a(7a) isolates FR1 amino acids 1-317.  
 XX  
 KW Hepatitis C virus; subtype; polymerase chain reaction; amplification;  
 KW PCR; primer; probe; antibody; infection.  
 XX  
 OS Hepatitis C virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 128  
 FT Misc-difference 129 /label= Phe, Ser, Tyr, Cys  
 FT Misc-difference 129 /label= Gly  
 FT /note= "amino acid in this position is designated X in  
 FT the specification, but codon usage shows that the only  
 FT possible amino acid at this pos. is Gly"  
 FT Misc-difference 308 /label= Ile, Met  
 FT Misc-difference 310 /label= Ser, Pro, Thr, Ala  
 FT  
 FT  
 XX WO9613590-A2.  
 XX  
 XX 09-MAY-1996.  
 PD  
 XX 23-OCT-1995; 95WO-EF004155.  
 PF  
 XX 21-OCT-1994; 94EP-00870166.  
 PR

PR 28-JUN-1995; 95EP-00870076.  
 XX  
 PA (INNO-) INNOGENETICS NV.  
 XX  
 FI Maertens G, Stuyver L;  
 XX  
 DR WPI; 1996-251460/25.  
 DR N-PSDB; AAT27957.  
 XX  
 XX Hepatitis C virus polynucleic acid unique to unidentified sub-type -  
 PT used to develop probes and primers for new subtypes and vaccines to  
 PT prevent and treat infection.  
 XX  
 PS Claim 25; Fig 3; 150pp; English.  
 XX  
 CC The sequences AAR96526-R96578 represent novel sequences isolated from  
 CC hepatitis C virus subtypes different from subtypes 1a-c, 2a-d, 3a-f, 4a-  
 CC j, 5a and 6a. They esp. from the novel subtypes 1d-f, 2e-i, 2k, 2l, 3g,  
 CC 4k-m, 7a-c or types 9, 10 or 11. The sequences corresp. to the 5',  
 CC untranslated region (UR), the Core/E1, NS4 or NS5B regions of the genome.  
 CC This sequence represents amino acids 1-317 from the HCV types 9a and 7a  
 CC isolates FR1. The new HCV types were isolated from patients with chronic  
 CC HCV from the Benelux countries, France, Cameroon and Vietnam, because of  
 CC their aberrant reactivities. The RNA was extracted, cDNA synthesised and  
 CC PCR amplified, cloned and genotyped. The 5'UR, Core/E1 and NS5B regions  
 CC were sequenced either directly or partially and used to classify the new  
 CC viruses into (sub)types based on comparison with known sequences. The  
 CC sequences were used to generate the peptides AAR96424-R96524. The  
 CC sequences can also be used to synthesise probes and primers for the  
 CC detection of HCV in a sample. The polypeptides can be used to detect anti  
 CC -HCV antibodies, for HCV typing or to prevent HCV infections  
 XX  
 SQ Sequence 319 AA;  
 Query Match 93.5%; Score 977; DB 2; Length 319;  
 Best Local Similarity 93.7%; Pred. No. 1.2e-83;  
 Matches 179; Conservative 3; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 MSTLPKQKTKRNTNRPTDVKFPGGGQIVGGVYLLPRGPRLGVRATRTKTSERSQPRG 60  
 DB 1 MSTLPKQKTKRNTNRPTDVKFPGGGQIVGGVYLLPRGPRLGVRATRTKTSERSQPRG 60  
 QY 61 RROPIKARQPOGRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGPNDRPRRSNGLG 120  
 DB 61 RROPIKVRQPTGRSWQPGYPWPLYGNEGCGWAGWLLSPRGRPNWGPNDRPRRSNGLG 120  
 QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180  
 DB 121 KVIDTLTXXLADLMGYIPVVGAPLGGVAAALAHGVRAIEDGVNATGNLPGCSFSIFLLA 180  
 QY 181 LLSCLTTPASA 191  
 DB 181 LLSCLTTPASA 191  
 RESULT 12  
 AAR22137  
 ID AAR22137 standard; protein; 326 AA.  
 XX  
 AC AAR22137;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 07-JUL-1992 (first entry)  
 XX  
 XX HCV-Hc59 capsid and envelope proteins.  
 DE  
 XX Hepatitis C virus; non-A non-B virus; HCV-Hc59; antigen; vaccine; assay;  
 KW detection.  
 XX  
 OS Non-A.  
 OS non-B hepatitis virus.  
 XX

```

FH Key Location/Qualifiers
FT 1. .326
FT /label= NANBV_structural_proteins
FT 1. .120
FT /label= capsid
FT 1. .74
FT /label= pref._capsid_antigen
FT 1. .20
FT /label= pref._capsid_antigen
FT 2. .40
FT /label= pref._capsid_antigen
FT Misc-difference 3
FT /label= Ile
FT /notes= "or Asn according to Seq No 46 (AAR22154), see CC"
FT Peptide 21. .40
FT /label= pref._capsid_antigen
FT Peptide 69. .120
FT /label= pref._capsid_antigen
FT Peptide 121. .326
FT /label= envelope_protein
FT 121. .176
FT /label= pref._envelope_antigen
FT Misc-difference 321
FT /label= Asp
FT /notes= "or Asn according to Seq No 1 (AAR22154), see CC"
XX WO9203458-A.
XX 05-MAR-1992.
XX 23-AUG-1991; 91WO-US006037.
XX 25-AUG-1990; 90US-00573643.
XX 21-NOV-1990; 90US-00616369.
XX 21-AUG-1991; 91US-00748564.
XX (NYBL-) NEW YORK BLOO DCENT.
XX (PHAR-) PHARMA.
XX Zebedee S, Inchauspe G, Nasofe MS, Prince AM;
XX WPI; 1992-096821/12.
XX N-PSDB; AAQ22838.
XX Deoxyribonucleic acid sequence encoding non-A, non-B hepatitis virus -
XX obtd. Hutch C59 subgroup encoding polypeptide(s), useful as vaccines, and
XX immuno reactive ABS for diagnosis of virus.
XX Disclosure; Page 131-133; 225pp; English.
XX One Hutch strain (HCV-H) of NANBV, designated the Hutch C59 isolate (HCV-
XX Hc59) was propagated through passage in animals and the entire viral
XX genome was cloned and sequenced (see AAQ22871). The sequence represented
XX here comprises two amino acids, indicated in the features, which differ
XX from the sequence of AAR22154. The proteins and peptides (see features)
XX and antibodies against them are useful for the prepn. of vaccines and
XX inoculums against NANBV and in immunological assays for detection of
XX viral infection. (Updated on 25-MAR-2003 to correct PR field.) (Updated
XX on 25-MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to correct
XX OS field.)
XX Sequence 326 AA;
Query Match 93.5%; Score 977; DB 2; Length 326;
Best Local Similarity 92.1%; Pred. No. 1.2e-83;
Matches 176; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
QY 1 MSTLPKQKTKRNTNRRPTDVKFGGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60
DB 1 MSTLPKQKTKRNTNRRPTDVKFGGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60
QY 61 RROQIPKARQPGRHWAQPGYPWPVLYGNEGCGWGLLSPRGSRPHWGNPDRRRSRNLG 120

```

```

DB 61 RROQIPKARQPGRHWAQPGYPWPVLYGNEGCGWGLLSPRGSRPHWGNPDRRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTTPASA 191
DB 181 LLSCLTTTPASA 191
RESULT 13
AAR38278
ID AAR38278 standard; protein; 733 AA.
XX AAR38278;
DT 21-OCT-1993 (first entry)
XX NANB hepatitis virus HC-OM gene polypeptide P-733-1.
DE Non-A, non-B; virus; polymerase chain reaction; detection; sensitive;
KW specific; HCV; NANBH.
XX Non-A.
OS non-B hepatitis virus.
XX JP05091884-A.
PD 16-APR-1993.
XX 10-APR-1991; 91JP-00196175.
XX 12-JUN-1990; 90JP-00153401.
XX 08-NOV-1990; 90JP-00304405.
XX (NAKA/) NAKAMURA T.
XX WPI; 1993-199637/25.
XX N-PSDB; AAQ43888.
XX Antigen related to non-A and non-B hepatitis virus - comprises non-
XX translation region comprising 340 - 341 mols. of nucleotides, non-
XX translation region comprising 1885 - 2551 mols. of nucleotides including
XX region 1,149 and, etc.
XX Claim 14; Page 32-35; 73pp; Japanese.
XX The sequence is that of NANB hepatitis virus HC-OM gene polypeptide P-733
XX -1. It may be used in a system for detecting NANB hepatitis. This
XX method is highly specific and sensitive, and can detect NANB hepatitis
XX virus which could not be detected by conventional methods
XX Sequence 733 AA;
Query Match 93.5%; Score 977; DB 2; Length 733;
Best Local Similarity 92.1%; Pred. No. 3e-83;
Matches 176; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
QY 1 MSTLPKQKTKRNTNRRPTDVKFGGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60
DB 1 MSTLPKQKTKRNTNRRPTDVKFGGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60
QY 61 RROQIPKARQPGRHWAQPGYPWPVLYGNEGCGWGLLSPRGSRPHWGNPDRRRSRNLG 120
DB 61 RROQIPKARQPGRHWAQPGYPWPVLYGNEGCGWGLLSPRGSRPHWGNPDRRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTTPASA 191

```

```

Db      181 LLSCLTVPASA 191

RESULT 14
ID AAR24440
XX AAR24440 standard; protein; 2894 AA.
XX AAR24440;
XX 25-MAR-2003 (revised)
DT 02-DEC-1992 (first entry)
XX Composite HCV HC-J1/CDC/CHI protein.
XX Hepatitis C virus; peptides; antibodies; ELISA.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..20
FT Peptide /label= 1
FT Peptide 7..26
FT Peptide /label= 2
FT Peptide 13..32
FT Peptide /label= 3
FT Peptide 37..56
FT Peptide /label= 4
FT Peptide 49..68
FT Peptide /label= 5
FT Peptide 61..80
FT Peptide /label= 6
FT Peptide 73..92
FT Peptide /label= 7
FT Peptide 1688..1707
FT Peptide /label= 8
FT Peptide 1694..1713
FT Peptide /label= 9
FT Peptide 1706..1725
FT Peptide /label= 10
FT Peptide 1712..1731
FT Peptide /label= 11
FT Peptide 1718..1737
FT Peptide /label= 12
FT Peptide 1724..1743
FT Peptide /label= 13
FT Peptide 1730..1749
FT Peptide /label= 14
FT Peptide 2263..2282
FT Peptide /label= 15
FT Peptide 2275..2294
FT Peptide /label= 16
FT Peptide 2287..2306
FT Peptide /label= 17
FT Peptide 2299..2318
FT Peptide /label= 18
FT Peptide 2311..2330
FT Peptide /label= 19
XX EP489968-A1.
PN
XX
XX 17-JUN-1992.
XX
XX 14-DEC-1990; 90EP-00124241.
XX
XX 14-DEC-1990; 90EP-00124241.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Deleys RJ, Pollet D, Maertens G, Van Heuverswyn H;
XX WPI; 1992-201383/25.
XX
XX New synthetic peptides for detecting antibodies to hepatitis C virus -

```

---

```

PT useful in e.g. ELISA assays, and for detection of HCV antigens or as
PT immunogens.
XX
PS Disclosure; Fig 1; 32pp; English.
XX
CC RNA viruses frequently exhibit a high rate of spontaneous mutation, thus
CC a virus is considered to be the same of equiv. to HCV if it exhibits a
CC global homology of more than 70 percent with the HCV HC- J1/CDC/CHI
CC composite sequence. The peptide fragments of this DNA sequence indicated
CC in the features table can immunologically mimic proteins encoded by HCV.
CC Additional amino acids or chemical gps. may be added to either end of the
CC peptides for the purpose of creating a linker arm for attachment to a
CC carrier. The peptides can be used for the detection of antibodies
CC specific for HCV. They may be used in the form of kits, opt. with
CC reagents such as staphylococcal protein A, streptococcal protein G,
CC avidin or streptavidin. The peptides may also be used as immunogens for
CC raising antibodies. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 2894 AA;
Query Match 93.5%; Score 977; DB 2; Length 2894;
Best Local Similarity 92.1%; Pred. No. 1.4e-82;
Matches 176; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
QY 1 MSTLPKPKORTKRNTRRPDVKPFPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
DB 1 MSTLPKPKORTKRNTRRPDVKPFPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
QY 61 RRQPIPKARQPGQRHWAQPGYPWPPLYGNEGCGWAGLLSPRGSRPHMGPNDRRRSRNLG 120
DB 61 RRQPIPKVRPEGRTWAPQPGYPWPPLYGNEGCGWAGLLSPRGSRPHMGPNDRRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVCAPLGCVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVVCAPLGCVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
DB 181 LLSCLTVPASA 191
RESULT 15
AAR70230
ID AAR70230 standard; protein; 2894 AA.
XX
XX AAR70230;
XX
XX 25-MAR-2003 (revised)
DT 07-NOV-1995 (first entry)
XX
XX Composite hepatitis C virus (HC-J1/CDC/CHI).
XX
XX Composite hepatitis C virus; HC-J1/CDC/CHI; HCV; non-A non-B;
XX synthetic antigens; blood screening.
XX
XX Hepatitis C virus.
XX
XX EP644202-A1.
PN
XX
XX 22-MAR-1995.
XX
XX 14-DEC-1990; 94EP-00108611.
XX
XX 14-DEC-1990; 90EP-00124241.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Deleys RJ, Pollet D, Maertens G, Van Heuverswyn H;
XX WPI; 1995-116946/16.
XX
XX Synthetic antigens for the detection of hepatitis C virus antibodies -
XX comprise portions of the HCV peptide sequence, for use in screening blood
PT

```



PT and blood products.

XX PS Disclosure; Fig 1; Sipp; English.

XX AAR70230 is the composite hepatitis C virus (HC-J1/CDC/CHI) protein from  
CC which the synthetic HCV antigens described in AAR70210-R70229 were  
CC derived. These synthetic antigens can be used to screen blood, or blood  
CC products for the presence of HCV, they can also be used in various specific  
CC assays for the detection of HCV antibodies, and antigens, or as  
CC immunogens. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-  
CC MAR-2003 to correct PF field.)

XX SQ Sequence 2894 AA;

Query Match 93.5%; Score 977; DB 2; Length 2894;  
Best Local Similarity 92.1%; Pred. NO. 1.4e-82;  
Matches 176; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
QY 1 MSTLPKQKTKNTNRRPTDVKFPGGGQIVGGVILLPRGPRGLGVRATRKTSERSQPRG 60  
DB 1 MSTLPKQKTKNTNRRPTDVKFPGGGQIVGGVILLPRGPRGLGVRATRKTSERSQPRG 60  
QY 61 RQPIPKARQPOGRHWAQPGYPWPLYNGEGCGWAGWLLSPGSRPHWGPNDRRSRLG 120  
DB 61 RQPIPKVRPEGRTWAGPYWPLYNGEGCGWAGWLLSPGSRPSWGPTDRRSRLG 120  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
QY 181 LLSCLTTPASA 191  
DB 181 LLSCLTVPASA 191

Search completed: July 15, 2005, 09:55:32  
Job time : 73 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2005, 09:50:42 ; Search time 22 Seconds  
(without alignments)  
648.090 Million cell updates/sec

Title: US-09-084-691B-206  
Perfect score: 1045  
Sequence: 1 MSTLPKPKRNTKRNRRPT.....CSFSIFLLALLSCLTTPASA 191

Scoring table: BLASTUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS-COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1045	100.0	191	2	US-08-290-665A-206
2	1045	100.0	191	5	PCT-US95-10398-206
3	1045	100.0	319	4	US-08-635-886C-228
4	1045	100.0	319	4	US-08-974-690C-228
5	996	95.3	191	2	US-08-290-665A-191
6	996	95.3	191	5	PCT-US95-10398-191
7	982	94.0	191	2	US-08-290-665A-192
8	982	94.0	191	2	US-08-290-665A-197
9	982	94.0	191	5	PCT-US95-10398-192
10	982	94.0	191	5	PCT-US95-10398-193
11	981	93.9	191	2	US-08-290-665A-172
12	981	93.9	191	5	PCT-US95-10398-172
13	980	93.8	191	2	US-08-290-665A-197
14	980	93.8	191	5	PCT-US95-10398-197
15	980	93.8	319	4	US-08-635-886C-217
16	980	93.8	319	4	US-08-974-690C-217
17	978	93.6	191	2	US-08-290-665A-195
18	978	93.6	191	5	PCT-US95-10398-195
19	977	93.5	191	2	US-08-290-665A-196
20	977	93.5	191	5	PCT-US95-10398-196
21	977	93.5	319	3	US-08-836-075A-42
22	977	93.5	319	4	US-08-635-886C-229
23	977	93.5	319	4	US-08-974-690C-229
24	977	93.5	450	4	US-08-635-886C-181
25	977	93.5	450	4	US-08-974-690C-181
26	977	93.5	2894	2	US-08-466-975A-23
27	977	93.5	2894	2	US-08-391-671A-23

28	977	93.5	2894	3	US-08-467-902A-23	Sequence 23, Appl
29	977	93.5	2894	3	US-09-275-265-23	Sequence 23, Appl
30	977	93.5	2894	4	US-09-941-611-23	Sequence 23, Appl
31	976	93.4	191	2	US-08-290-665A-156	Sequence 156, App
32	976	93.4	191	2	US-08-290-665A-157	Sequence 157, App
33	976	93.4	191	2	US-08-290-665A-158	Sequence 158, App
34	976	93.4	191	2	US-08-290-665A-159	Sequence 159, App
35	976	93.4	191	2	US-08-290-665A-160	Sequence 160, App
36	976	93.4	191	3	US-08-380-160-3	Sequence 3, Appl
37	976	93.4	191	5	PCT-US95-10398-156	Sequence 156, App
38	976	93.4	191	5	PCT-US95-10398-157	Sequence 157, App
39	976	93.4	191	5	PCT-US95-10398-158	Sequence 158, App
40	976	93.4	191	5	PCT-US95-10398-159	Sequence 159, App
41	976	93.4	191	5	PCT-US95-10398-160	Sequence 160, App
42	976	93.4	450	4	US-08-635-886C-179	Sequence 179, App
43	976	93.4	450	4	US-08-635-886C-180	Sequence 180, App
44	976	93.4	450	4	US-08-974-690C-179	Sequence 179, App
45	976	93.4	450	4	US-08-974-690C-180	Sequence 180, App

## ALIGNMENTS

RESULT 1  
US-08-290-665A-206  
; Sequence 206, Application US/08290665A  
; Patent No. 5882852  
; GENERAL INFORMATION:  
; APPLICANT: BUKH, J., MILLER, R.H. AND  
; APPLICANT: PURCELL, R.H.  
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED  
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND  
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS  
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE  
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES  
; NUMBER OF SEQUENCES: 263  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/290.665A  
; FILING DATE: 15-AUG-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: RICHARD W. BORK  
; REGISTRATION NUMBER: 36,459  
; REFERENCE/DOCKET NUMBER: 2026-4116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; TELEX: 421792  
; INFORMATION FOR SEQ ID NO: 206:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 191 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; ORIGINAL SOURCE:  
; ORGANISM: hom sapiens  
; INDIVIDUAL ISOLATE: HK2  
; US-08-290-665A-206

Query Match 100.0%; Score 1045; DB 2; Length 191;  
Best Local Similarity 100.0%; Pred. No. 9.9e-97;



APPLICANT: MAERTENS, Geert  
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
FILE REFERENCE: 2551-94  
CURRENT APPLICATION NUMBER: US/08/974,690C  
PRIOR FILING DATE: 1997-11-19  
PRIOR APPLICATION NUMBER: PCT/EP94/03555  
PRIOR FILING DATE: 1994-10-28  
PRIOR APPLICATION NUMBER: EP 93402718.6  
PRIOR FILING DATE: 1993-11-04  
NUMBER OF SEQ ID NOS: 286  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 228  
LENGTH: 319  
TYPE: PRT  
ORGANISM: hepatitis C virus  
US-08-974-690C-228

Query Match 100.0%; Score 1045; DB 4; Length 319;  
Best Local Similarity 100.0%; Pred. No. 1.9e-96;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTLPKQKTKRNTNRRPTDVKFGGGQIVGGVLLPRGRLGVRATRKTSERSQPRG 60  
DB 1 MSTLPKQKTKRNTNRRPTDVKFGGGQIVGGVLLPRGRLGVRATRKTSERSQPRG 60

QY 61 RQPIPKARQPGQRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGPNDRRRSRNLG 120  
DB 61 RQPIPKARQPGQRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGPNDRRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191  
DB 181 LLSCLTTPASA 191

RESULT 5  
US-08-290-665A-191  
Sequence 191, Application US/08290665A  
Patent No. 5882852  
GENERAL INFORMATION:  
APPLICANT: BUKH, J., MILLER, R.H. AND  
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED  
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND  
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS  
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE  
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES  
NUMBER OF SEQUENCES: 263  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/290,665A  
FILING DATE: 15-AUG-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: RICHARD W. BORK  
REGISTRATION NUMBER: 36,459  
REFERENCE/DOCKET NUMBER: 2026-4116  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 191:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 191 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
ORIGINAL SOURCE:  
ORGANISM: homosapiens  
INDIVIDUAL ISOLATE: Z4  
US-08-290-665A-191

Query Match 95.3%; Score 996; DB 2; Length 191;  
Best Local Similarity 94.8%; Pred. No. 7.9e-92;  
Matches 181; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 MSTLPKQKTKRNTNRRPTDVKFGGGQIVGGVLLPRGRLGVRATRKTSERSQPRG 60  
DB 1 MSTLPKQKTKRNTNRRPTDVKFGGGQIVGGVLLPRGRLGVRATRKTSERSQPRG 60

QY 61 RQPIPKARQPGQRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGPNDRRRSRNLG 120  
DB 61 RQPIPKARQPGQRHWAQPGYPWPLYGNEGCGWAGWLLSPRGRPHWGPNDRRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191  
DB 181 LLSCLTTPASA 191

RESULT 6  
PCT-US95-10398-191  
Sequence 191, Application PC/TUS9510398  
GENERAL INFORMATION:  
APPLICANT: BUKH, J., MILLER, R.H. AND  
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED  
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND  
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS  
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE  
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES  
NUMBER OF SEQUENCES: 263  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/10398  
FILING DATE: 15-AUG-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/086,428  
FILING DATE: 29 JUNE 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/290/665  
FILING DATE: 15 AUGUST 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: RICHARD W. BORK  
REGISTRATION NUMBER: 36,459  
REFERENCE/DOCKET NUMBER: 2026-4116

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 191:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: 24
; PCT-US95-10398-191

Query Match 95.3%; Score 996; DB 5; Length 191;
Best Local Similarity 94.8%; Pred. No. 7.9e-92;
Matches 181; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MSTLPKPKQKRNTRRPTDVKFGGQIVGGVYLLPRRGRLGVRAATKTSERSQPRG 60
Db 1 MSTNPKPKQKRNTRRPMDDVKFGGQIVGGVYLLPRRGRLGVRAATKTSERSQPRG 60

Qy 61 RROPIPKARQPGCRHWAQPGYPWPPLYGNEGCGWAGWLLSPRGRPHMGPNDRRSRLG 120
Db 61 RROPIPKARQPGCRHWAQPGYPWPPLYGNEGCGWAGWLLSPRGRPHMGPNDRRSRLG 120

Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 7
US-08-290-665A-192
; Sequence 192, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792

; INFORMATION FOR SEQ ID NO: 192:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: Z8
; US-08-290-665A-192

Query Match 94.0%; Score 982; DB 2; Length 191;
Best Local Similarity 93.7%; Pred. No. 2e-90;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MSTLPKPKQKRNTRRPTDVKFGGQIVGGVYLLPRRGRLGVRAATKTSERSQPRG 60
Db 1 MSTNPKPKQKRNTRRPMDDVKFGGQIVGGVYLLPRRGRLGVRAATKTSERSQPRG 60

Qy 61 RROPIPKARQPGCRHWAQPGYPWPPLYGNEGCGWAGWLLSPRGRPHMGPNDRRSRLG 120
Db 61 RROPIPKARQPGCRHWAQPGYPWPPLYGNEGCGWAGWLLSPRGRPHMGPNDRRSRLG 120

Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 8
US-08-290-665A-193
; Sequence 193, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 193:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid

```

```
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: Z1
US-08-290-665A-193

Query Match          94.0%; Score 982; DB 2; Length 191;
Best Local Similarity 93.7%; Pred. No. 2e-90;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSTLPKQKTKNTNRRPTDVKFPGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60
   |||||
Db 1 MSTNPKQKTKNTNRRPMDVKFPGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60
   |||||

QY 61 RROPIPKARQPCGRHWAQPGYWPPLYGNEGCGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
   |||||
Db 61 RROPIPKARSEGRSWAQPGYWPPLYGNEGCGWAGWLLSPRGRSPWGPNDRRRSRNLG 120
   |||||

QY 121 KVDTLTGCFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
   |||||
Db 121 KVDTLTGCFADLMGYIPLVGAPVGGVARALAHGVRAVEDGINYATGNLPGCSFSIFLLA 180
   |||||

QY 181 LLSCLTTPASA 191
   |||||
Db 181 LLSCLTTPASA 191
   |||||
```

```
RESULT 9
PCT-US95-10398-192
; Sequence 192, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10398
; FILING DATE: 15-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29 JUNE 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290/665
; FILING DATE: 15 AUGUST 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 192:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
```

```
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: Z8
PCT-US95-10398-192

Query Match          94.0%; Score 982; DB 5; Length 191;
Best Local Similarity 93.7%; Pred. No. 2e-90;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSTLPKQKTKNTNRRPTDVKFPGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60
   |||||
Db 1 MSTNPKQKTKNTNRRPMDVKFPGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60
   |||||

QY 61 RROPIPKARQPCGRHWAQPGYWPPLYGNEGCGWAGWLLSPRGRPHWGPNDRRRSRNLG 120
   |||||
Db 61 RROPIPKARSEGRSWAQPGYWPPLYGNEGCGWAGWLLSPRGRSPWGPNDRRRSRNLG 120
   |||||

QY 121 KVDTLTGCFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
   |||||
Db 121 KVDTLTGCFADLMGYIPLVGAPVGGVARALAHGVRAVEDGINYATGNLPGCSFSIFLLA 180
   |||||

QY 181 LLSCLTTPASA 191
   |||||
Db 181 LLSCLTTPASA 191
   |||||
```

```
RESULT 10
PCT-US95-10398-193
; Sequence 193, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10398
; FILING DATE: 15-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29 JUNE 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290/665
; FILING DATE: 15 AUGUST 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 193:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: hominids
; INDIVIDUAL ISOLATE: 21
PCT-US95-10398-193

Query Match          94.0%; Score 982; DB 5; Length 191;
Best Local Similarity 93.7%; Pred. No. 2e-90;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKRNTRRPTDVKFGGQIVGGVLLPRGRLGVRATKTSRSQPRG 60
DB 1 MSTNPKPKQKRNTRRPTDVKFGGQIVGGVLLPRGRLGVRATKTSRSQPRG 60
QY 61 RQPIPKARQPGGRHWAQPGYPWPLYGNEGCGWAGWLLSPRSGPHMGPNDRSRNLG 120
DB 61 RQPIPKARQPGGRHWAQPGYPWPLYGNEGCGWAGWLLSPRSGPHMGPNDRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 11
US-08-290-665A-172
; Sequence 172, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 172:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown

; ORIGINAL SOURCE:
; ORGANISM: hominids
; INDIVIDUAL ISOLATE: 21
US-08-290-665A-172

Query Match          93.9%; Score 981; DB 2; Length 191;
Best Local Similarity 93.7%; Pred. No. 2.5e-90;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKRNTRRPTDVKFGGQIVGGVLLPRGRLGVRATKTSRSQPRG 60
DB 1 MSTNPKPKQKRNTRRPTDVKFGGQIVGGVLLPRGRLGVRATKTSRSQPRG 60
QY 61 RQPIPKARQPGGRHWAQPGYPWPLYGNEGCGWAGWLLSPRSGPHMGPNDRSRNLG 120
DB 61 RQPIPKARQPGGRHWAQPGYPWPLYGNEGCGWAGWLLSPRSGPHMGPNDRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 12
PCT-US95-10398-172
; Sequence 172, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10398
; FILING DATE: 15-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29 JUNE 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290/665
; FILING DATE: 15 AUGUST 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 172:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
```



```

; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: HK3
PCT-US95-10398-172

Query Match      93.9%; Score 981; DB 5; Length 191;
Best Local Similarity 93.7%; Pred. No. 2.5e-90;
Matches 179; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKNTNRRPTDVKFPGGGQIVGGVYLLPRGPRGLGVRATRKTSERSQPRG 60
Db 1 MSTNPKPQKTKNTNRRPQDVKFPFGGGQIVGGVYLLPRGPRGLGVRATRKTSERSQPRG 60

QY 61 RROPIPKARQPCQRRHWAQPGYWPVLYGNEGCGWAGWLLSPRSGRPHWGPNDRRRSRNLG 120
Db 61 RROPIPKARQPCGRTHWAQPGYWPVLYGNEGCGWAGWLLSPRSGRPHWGPNDRRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 13
US-08-290-665A-197
; Sequence 197, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 197:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: DK13

US-08-290-665A-197
Query Match      93.8%; Score 980; DB 2; Length 191;
Best Local Similarity 93.7%; Pred. No. 3.1e-90;
Matches 179; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKNTNRRPTDVKFPGGGQIVGGVYLLPRGPRGLGVRATRKTSERSQPRG 60
Db 1 MSTNPKPQKTKNTNRRPMDVKFPFGGGQIVGGVYLLPRGPRGLGVRATRKTSERSQPRG 60

QY 61 RROPIPKARQPCQRRHWAQPGYWPVLYGNEGCGWAGWLLSPRSGRPHWGPNDRRRSRNLG 120
Db 61 RROPIPKARQPCGRTHWAQPGYWPVLYGNEGCGWAGWLLSPRSGRPHWGPNDRRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 14
PCT-US95-10398-197
; Sequence 197, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10398
; FILING DATE: 15-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29 JUNE 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290/665
; FILING DATE: 15 AUGUST 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 197:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homosapiens

```

```

; INDIVIDUAL ISOLATE: DK13
PCT-US95-10398-197

Query Match          93.8%; Score 980; DB 5; Length 191;
Best Local Similarity 93.7%; Pred. No. 3.1e-90;
Matches 179; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRRPTDVKFGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
Db 1 MSTNPKPKQKTKRNTNRRPMDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60

QY 61 RRQPIPKARQPGQRHWAQPGYWPPLYGNEGCWAGWLLSPRGRPHWGPNDPRRRSRNLG 120
Db 61 RRQPIPKARQPGQRHWAQPGYWPPLYGNEGCWAGWLLSPRGRPSWGPNDPRRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPVGGVARALAHGVRLLEDGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
Db 181 LLSCLTVPASA 191

```

```

RESULT 15
US-08-635-886C-217
; Sequence 217, Application US/08635886C
; Patent No. 6555114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635,886C
; PRIOR FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 217
; LENGTH: 319
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-217

```

```

Query Match          93.8%; Score 980; DB 4; Length 319;
Best Local Similarity 93.7%; Pred. No. 5.9e-90;
Matches 179; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRRPTDVKFGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
Db 1 MSTNPKPKQKTKRNTNRRPMDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60

QY 61 RRQPIPKARQPGQRHWAQPGYWPPLYGNEGCWAGWLLSPRGRPHWGPNDPRRRSRNLG 120
Db 61 RRQPIPKARQPGQRHWAQPGYWPPLYGNEGCWAGWLLSPRGRPSWGPNDPRRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPVGGVARALAHGVRLLEDGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
Db 181 LLSCLTVPASA 191

```

Search completed: July 15, 2005, 09:59:02  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen, Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2005, 09:58:44 ; Search time 164 Seconds  
(without alignments)  
450.590 Million cell updates/sec

Title: US-09-084-691B-206

Perfect score: 1045  
Sequence: 1 MSTLPKPKQKTKRNTNRRPT.....CSFSIFLLALLSCLTTPASA 191

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1729182 seqs, 386893608 residues

Total number of hits satisfying chosen parameters: 1729182

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1045	100.0	319	15	US-10-651-165-228
2	980	93.8	319	15	US-10-651-165-217
3	977	93.5	319	9	US-09-851-138-42
4	977	93.5	319	15	US-10-651-165-229
5	977	93.5	450	15	US-10-651-165-181
6	977	93.5	2894	9	US-09-941-611-23
7	977	93.5	2894	14	US-10-044-995-23
8	977	93.5	2894	17	US-10-822-871-23
9	976	93.4	249	15	US-10-365-620-34
10	976	93.4	249	17	US-10-912-969-56
11	976	93.4	450	15	US-10-651-165-179

12	976	93.4	450	15	US-10-651-165-180	Sequence 180, App
13	976	93.4	473	15	US-10-365-620-56	Sequence 56, Appl
14	976	93.4	473	17	US-10-912-969-58	Sequence 58, Appl
15	976	93.4	473	17	US-10-913-171-39	Sequence 39, Appl
16	976	93.4	3011	9	US-09-742-659-4	Sequence 4, Appl
17	976	93.4	3011	9	US-09-952-572-9	Sequence 9, Appl
18	976	93.4	3011	9	US-09-929-955-1	Sequence 1, Appl
19	976	93.4	3011	9	US-09-747-419-20	Sequence 20, Appl
20	976	93.4	3011	10	US-09-891-894-3	Sequence 3, Appl
21	976	93.4	3011	13	US-10-104-966-1	Sequence 1, Appl
22	976	93.4	3011	14	US-10-259-275-20	Sequence 20, Appl
23	976	93.4	3011	14	US-10-184-150-3	Sequence 3, Appl
24	976	93.4	3011	15	US-10-328-997-3	Sequence 3, Appl
25	976	93.4	3011	15	US-10-189-359-14	Sequence 14, Appl
26	976	93.4	3011	15	US-10-296-734-406	Sequence 406, App
27	976	93.4	3011	15	US-10-719-619-1	Sequence 1, Appl
28	976	93.4	3011	16	US-10-817-591-1	Sequence 1, Appl
29	976	93.4	3012	9	US-09-238-076-2	Sequence 2, Appl
30	976	93.4	3012	10	US-09-995-937-2	Sequence 2, Appl
31	976	93.4	3012	10	US-09-917-563-2	Sequence 2, Appl
32	972	93.0	3011	9	US-09-238-076-20	Sequence 20, Appl
33	972	93.0	3011	10	US-09-995-937-20	Sequence 20, Appl
34	972	93.0	3011	10	US-09-917-563-20	Sequence 20, Appl
35	970	92.8	450	15	US-10-651-165-190	Sequence 190, App
36	969	92.7	450	15	US-10-651-165-189	Sequence 189, App
37	969	92.7	3010	15	US-10-467-000-1	Sequence 1, Appl
38	968	92.6	191	18	US-10-770-117-4	Sequence 4, Appl
39	968	92.6	319	15	US-10-651-165-219	Sequence 219, App
40	968	92.6	3011	9	US-09-916-359-2	Sequence 2, Appl
41	968	92.6	3011	16	US-10-445-724-2	Sequence 2, Appl
42	967	92.5	191	10	US-09-194-949-3	Sequence 3, Appl
43	967	92.5	191	17	US-10-664-391-3	Sequence 3, Appl
44	967	92.5	191	18	US-10-770-117-2	Sequence 2, Appl
45	965	92.3	450	15	US-10-651-165-187	Sequence 187, App

#### ALIGNMENTS

##### RESULT 1

US-10-651-165-228  
; Sequence 228, Application US/10651165  
; Publication No. US20040047877A1  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, Geert  
; APPLICANT: DELEYS, Robert  
; APPLICANT: MAERTENS, Geert  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
; TITLE OF INVENTION: VIRUS  
; FILE REFERENCE: 2551-94  
; CURRENT APPLICATION NUMBER: US/10/651,165  
; CURRENT FILING DATE: 2003-09-02  
; PRIOR APPLICATION NUMBER: US/08/974,690C  
; PRIOR FILING DATE: 1997-11-19  
; PRIOR APPLICATION NUMBER: PCT/EP94/03555  
; PRIOR FILING DATE: 1994-10-28  
; PRIOR APPLICATION NUMBER: EP 93402718.6  
; PRIOR FILING DATE: 1993-11-04  
; NUMBER OF SEQ ID NOS: 286  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 228  
; LENGTH: 319  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
US-10-651-165-228

Query Match 100.0%; Score 1045; DB 15; Length 319;  
Best Local Similarity 100.0%; Pred. No. 1.5e-83;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSTLPKPKQKTKRNTNRRPTDKFPGGQIVGGVYLLPRRGRPLGVTRATKTSRSQPRG 60  
Db 1 MSTLPKPKQKTKRNTNRRPTDKFPGGQIVGGVYLLPRRGRPLGVTRATKTSRSQPRG 60

QY 61 RRQIPKARQPOGRHWAQPGYPWLYGNEGCGWAGWLLSPRGRPHWGPNDRPRRSRLG 120  
Db |||||||  
QY 61 RRQIPKARQPOGRHWAQPGYPWLYGNEGCGWAGWLLSPRGRPHWGPNDRPRRSRLG 120  
Db |||||||  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
Db |||||||  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
Db |||||||  
QY 181 LLSCLTTPASA 191  
Db |||||||  
QY 181 LLSCLTTPASA 191  
Db |||||||

## RESULT 2

US-10-651-165-217  
; Sequence 217, Application US/10651165  
; Publication No. US20040047877A1  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, Geert  
; APPLICANT: DELEYS, Robert  
; APPLICANT: MAERTENS, Geert  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
; FILE REFERENCE: 2551-94  
; CURRENT APPLICATION NUMBER: US/10/651,165  
; PRIOR FILING DATE: 2003-09-02  
; PRIOR APPLICATION NUMBER: US/08/974,690C  
; PRIOR FILING DATE: 1997-11-19  
; PRIOR APPLICATION NUMBER: PCT/EP94/03555  
; PRIOR FILING DATE: 1994-10-28  
; PRIOR APPLICATION NUMBER: EP 93402718.6  
; PRIOR FILING DATE: 1993-11-04  
; NUMBER OF SEQ ID NOS: 286  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 217  
; LENGTH: 319  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
US-10-651-165-217

Query Match 93.8%; Score 980; DB 15; Length 319;  
Best Local Similarity 93.7%; Pred. No. 7.6e-78;  
Matches 179; Conservative 4; Mismatches 8; Indels 0; Gaps 0;  
QY 1 MSTLPKQKTKRNTNRPTDVKFGGGQIVGGVYLLPRRGLGVRAIATKTSERSQPRG 60  
Db |||||||  
QY 1 MSTNPKPQKTKRNTNRPTDVKFGGGQIVGGVYLLPRRGLGVRAIATKTSERSQPRG 60  
Db |||||||  
QY 61 RRQIPKARQPOGRHWAQPGYPWLYGNEGCGWAGWLLSPRGRPHWGPNDRPRRSRLG 120  
Db |||||||  
QY 61 RRQIPKARQPOGRHWAQPGYPWLYGNEGCGWAGWLLSPRGRPHWGPNDRPRRSRLG 120  
Db |||||||  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
Db |||||||  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
Db |||||||  
QY 181 LLSCLTTPASA 191  
Db |||||||  
QY 181 LLSCLTTPASA 191  
Db |||||||

## RESULT 3

US-09-851-138-42  
; Sequence 42, Application US/09851138  
; Publication No. US20020183508A1  
; GENERAL INFORMATION:  
; APPLICANT: MAERTENS, GEERT  
; TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES  
; AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC  
; AGENTS  
; NUMBER OF SEQUENCES: 207

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ARNOLD, WHITE & DURKEE  
; STREET: P. O. BOX 4433  
; CITY: HOUSTON  
; STATE: TEXAS  
; COUNTRY: USA  
; ZIP: 77210-4433  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Microsoft Word 6.0 / ASCII text output  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/851,138  
; FILING DATE: 09-May-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/836,075  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: EP 94870166.9  
; FILING DATE: 21 Oct 1994  
; APPLICATION NUMBER: EP 95870076.7  
; FILING DATE: 28 Jun 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KAMMERER, PATRICIA A.  
; REGISTRATION NUMBER: 29,775  
; REFERENCE/DOCKET NUMBER: INNS:004  
; INFORMATION FOR SEQ ID NO: 42:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 319 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 42:  
US-09-851-138-42

Query Match 93.5%; Score 977; DB 9; Length 319;  
Best Local Similarity 93.7%; Pred. No. 1.4e-77;  
Matches 179; Conservative 3; Mismatches 9; Indels 0; Gaps 0;  
QY 1 MSTLPKQKTKRNTNRPTDVKFGGGQIVGGVYLLPRRGLGVRAIATKTSERSQPRG 60  
Db |||||||  
QY 1 MSTLPKQKTKRNTNRPTDVKFGGGQIVGGVYLLPRRGLGVRAIATKTSERSQPRG 60  
Db |||||||  
QY 61 RRQIPKARQPOGRHWAQPGYPWLYGNEGCGWAGWLLSPRGRPHWGPNDRPRRSRLG 120  
Db |||||||  
QY 61 RRQIPKARQPOGRHWAQPGYPWLYGNEGCGWAGWLLSPRGRPHWGPNDRPRRSRLG 120  
Db |||||||  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
Db |||||||  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
Db |||||||  
QY 181 LLSCLTTPASA 191  
Db |||||||  
QY 181 LLSCLTTPASA 191  
Db |||||||

## RESULT 4

US-10-651-165-229  
; Sequence 229, Application US/10651165  
; Publication No. US20040047877A1  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, Geert  
; APPLICANT: DELEYS, Robert  
; APPLICANT: MAERTENS, Geert  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
; FILE REFERENCE: 2551-94  
; CURRENT APPLICATION NUMBER: US/10/651,165  
; CURRENT FILING DATE: 2003-09-02  
; PRIOR APPLICATION NUMBER: US/08/974,690C  
; PRIOR FILING DATE: 1997-11-19  
; PRIOR APPLICATION NUMBER: PCT/EP94/03555  
; PRIOR FILING DATE: 1994-10-28

```

; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 229
; LENGTH: 319
; TYPE: PRT
; ORGANISM: hepatitis C virus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (128)..(129)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (308)..(308)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (310)..(310)
; OTHER INFORMATION: Xaa is any amino acid
; US-10-651-165-229

Query Match          93.5%; Score 977; DB 15; Length 319;
Best Local Similarity 93.7%; Pred. No. 1.4e-77;
Matches 179; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRRPTDVKFGGQIVGGVYLLPRRGLGVRAVTRKTSERSQPRG 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MSTLPKPKQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRAVTRKTSERSQPRG 60

QY 61 RROPIPKARQPGGRHWAQPGYPWPPLYGNEGCGWAGWLLSPRGRPHWGNPDRRSRLNG 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 RROPIPKVRPEGRTWAGPGYPWPPLYGNEGCGWAGWLLSPRGRSPSWGPTDPRRSRLNG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 KVIDTLTCGFADLMGYIPLVCGAPLGGVAAALAHGVRAIEDGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 LLSCLTTPASA 191

RESULT 6
US-09-941-611-23
; Sequence 23, Application US/09941611
; Patent No. US20020106640A1
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; POLLET, DIRK
; MAERTENS, GEERT
; VAN HEUVERSUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/941,611
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/391,671
; FILING DATE: 1995-02-21
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; APPLICATION NUMBER: EP 90124241.2
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2894 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-941-611-23

Query Match          93.5%; Score 977; DB 9; Length 2894;
US-10-651-165-181

Query Match          93.5%; Score 977; DB 15; Length 450;
Best Local Similarity 92.1%; Pred. No. 2e-77;
Matches 176; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 229
; LENGTH: 319
; TYPE: PRT
; ORGANISM: hepatitis C virus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (128)..(129)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (308)..(308)
; OTHER INFORMATION: Xaa is any amino acid
; US-10-651-165-229

Query Match          93.5%; Score 977; DB 15; Length 319;
Best Local Similarity 93.7%; Pred. No. 1.4e-77;
Matches 179; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRRPTDVKFGGQIVGGVYLLPRRGLGVRAVTRKTSERSQPRG 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MSTLPKPKQKTKRNTNRRPMDVKFGGQIVGGVYLLPRRGLGVRAVTRKTSERSQPRG 60

QY 61 RROPIPKARQPGGRHWAQPGYPWPPLYGNEGCGWAGWLLSPRGRPHWGNPDRRSRLNG 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 RROPIPKVRPEGRTWAGPGYPWPPLYGNEGCGWAGWLLSPRGRSPSWGPTDPRRSRLNG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 KVIDTLTXLADLMGYIPVLGFLGGVAAALAHGVRAIEDGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 LLSCLTTPASA 191

RESULT 5
US-10-651-165-181
; Sequence 181, Application US/10651165
; Publication No. US20040047877A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 181
; LENGTH: 450
; TYPE: PRT
; ORGANISM: hepatitis C virus
; US-10-651-165-181
```

```

Best Local Similarity 92.1%; Pred: No. 1.5e-76;
Matches 176; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

1 MSTLPKQRTKRNTRRPTDVKFPGGQIVGGVYLLPRRGPRLCGVRA1TRKTSERS2PRG 60
  |||:|||||
1 MSTIPKQRTKRNTRRBPQDVKFPGGQIVGGVYLLPRRGPRLCGVRA1TRKTSERS2PRG 60
  |||:|||||
61 RRQPTPKARQPGQRHWAQPGYPWPLYGNEGCGWAGWLLS1PGSRPHWCPNDPRRRS2NLG 120
  |||:|||||
61 RRQPTPKVRRPEGR1TWAGPGYPWPLYGNEGCGWAGWLLS1PGSRSPSWGPTDPRRRS2NLG 120
  |||:|||||
121 KVIDTLTCGFADLMGYIIPVWGAPLGGVAAALAHGVRA1TEIDGINVATGNLPGCSFS2IFLLA 180
  |||:|||||
121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRV1LEDGVNATGNLPGCSFS2IFLLA 180
  |||:|||||
181 LLSCLTTPASA 191
  |||||
181 LLSCLTTPASA 191
  |||||

```

RESULT 7  
8-10-044-995-23  
Sequence 23, Application US/10044995  
Publication No. US20030049685A1  
GENERAL INFORMATION:  
APPLICANT: DELEYS, ROBERT J  
FOLLET, DIRK  
MAERTENS, GEERT  
VAN HEUVERSWUN, HUGO  
TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF  
ANTIBODIES TO HEPATITIS C VIRUS

```

; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-10-044-995-23

Query Match          93.5%; Score 977; DB 14; Length 2894;
Best Local Similarity 92.1%; Pred. No. 1.5e-76;
Matches 176; Conservative 6; Mismatches 9; Indels 0;

Qy 1 MSTLPKQPKTKENTNRRPTDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSSE
Db 1 MSTIPKQPKTKENTNRRPDQVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSSE

Qy 61 RROPDKARQPOGRHWAQGYGYPWLYGNEGCGWAGWLLSPRGSRPHWGNDPRRR
Db 61 RROPDKARQPOGRHWAQGYGYPWLYGNEGCGWAGWLLSPRGSRPHWGNDPRRR

Qy 121 KVIDTUTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSF
Db 121 KVIDTUTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSF

Qy 181 LLSCLTTPASA 191
Db 181 LLSCLTVPASA 191

RESULT 8
US-10-822-871-23
; Sequence 23, Application US/10822871
; Publication No. US20050003345A1
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; POLLET, DIRK
; MAERTENS, GEERT
; VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSER: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/822,871
; FILING DATE: 13-Apr-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/044,995
; FILING DATE: 15-Jan-2002
; APPLICATION NUMBER: 08/391,671
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/920,286
; FILING DATE: 14-OCT-1992
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; APPLICATION NUMBER: EP 90124241.2
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2894 amino acids

```

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-10-822-871-23

Query Match      93.5%; Score 977; DB 17; Length 2894;
Best Local Similarity 92.1%; Pred. No. 1.5e-76;
Matches 176; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 MSTLPKQKTKNTNRRPTDVKPGGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MSTIPKQKTKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 RROPIPKARQPGRRHWAQPGYWPVLYGNEGCWAGWLLSPRGRPHWGPNDRRRSRLG 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 RROPIPKARRPEGRTWAOQGYWPVLYGNEGCWAGWLLSPRGRPSWGTDPDRRSRLG 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRAIEDGVNATGNLPGCSFSIFLLA 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 LLSCLTTPASA 191
   |||||:|||||
Db 181 LLSCLTVPASA 191
   |||||:|||||

RESULT 9
US-10-365-620-54
; Sequence 54, Application US/10365620
; Publication No. US20040001853A1
; GENERAL INFORMATION:
; APPLICANT: George, Rajan
; APPLICANT: Tyrrell, Lorne
; TITLE OF INVENTION: Chimeric Antigens for Eliciting An Immune Response
; FILE REFERENCE: 656.0016
; CURRENT APPLICATION NUMBER: US/10/365,620
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US60/423,578
; PRIOR FILING DATE: 2003-11-05
; PRIOR APPLICATION NUMBER: 60/390,564
; PRIOR FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 54
; LENGTH: 249
; TYPE: PRT
; ORGANISM: ORF of HCV Core Protein
US-10-365-620-54

Query Match      93.4%; Score 976; DB 15; Length 249;
Best Local Similarity 92.7%; Pred. No. 1.3e-77;
Matches 177; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 MSTLPKQKTKNTNRRPTDVKPGGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 31 MSTNPKQKTKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 90
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 RROPIPKARQPGRRHWAQPGYWPVLYGNEGCWAGWLLSPRGRPHWGPNDRRRSRLG 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 91 RROPIPKARRPEGRTWAOQGYWPVLYGNEGCWAGWLLSPRGRPSWGTDPDRRSRLG 150
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 151 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRAIEDGVNATGNLPGCSFSIFLLA 210
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 LLSCLTTPASA 191
   |||||:|||||
Db 211 LLSCLTVPASA 221
   |||||:|||||

US-10-651-165-179
; Sequence 179, Application US/10651165
; Publication No. US2004004787A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
```

```

; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 179
; LENGTH: 450
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-651-165-179

```

	Query Match	93.4%	Score 976;	DB 15;	Length 450;
	Best Local Similarity	92.7%;	Pred. No. 2.5e-77;		
	Matches 177; Conservative	5;	Mismatches 9;	Indels 0;	Gaps 0;
Qy	1	MSTLPKPORTKTNTNRPRDVKFPPGGGQIVGGVYLLPRGRPLGVTRATKTTSRSPRG	60		
Db	1	MSINPFORTKTNTNRPRQDVKEPGGGQIVGGVYLLPRGRPLGVTRATKTTSRSPRG	60		
Qy	61	RROPIPKARPOQGRHWAQPYPWPPLYNVEGCWAGWILLSPRGSRPHWGPNDRRRSRNLG	120		
		:   :   :			
Db	61	RROPIPKARRPEGRTWAPGPYPWPPLYNVEGCWAGWILLSPRGSRPSWGPTDPRRSRNLG	120		
		:   :   :			
Qy	121	KVIDTLTCGFADLMGYTPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPCGCSFSIFLLA	180		
		:   :   :			
Db	121	KVIDTLTCGFADLMGYIPLVGAPLGGGAARLAHHGVRLVEDGVNYATGNLPCGCSFSIFLLA	180		
		:   :   :			
Qy	181	LLSCLTTPASA	191		
Db	181	LLSCLTVFASA	191		

```

RESULT 12
US-10-651-165-180
; Sequence 180, Application US/10651165
; Publication No. US20040047877A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
; CURRENT FILING DATE: 2003-09-02
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 180
; LENGTH: 450
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-651-165-180

```

Query Match	93.4%	Score 976;	DB 15;	Length 450;
Best Local Similarity	92.7%;	Pred. No. 2.5e-77;		
Matches 177;	Conservative 5;	Mismatches 9;	Indels 0;	Gaps 0;
QY	1	MSTLKPQKTKENTNRPRDVKFPGGGQIVGGVYLLPRGPRLGVRATKTKTSERS	PRG 60	
DB	1	MSTNPKPQKTKENTNRPRDVKFPGGGQIVGGVYLLPRGPRLGVRATKTKTSERS	PRG 60	
QY	61	RRQPIPKARQPOQRHWAQPGYPWPLYNCGCGWAGWLLSPRGSRPHWGPNDRPRRS	NLG 120	
DB	61	RRQPIPKARPEGRWAQPGYPWPLYNCGCGWAGWLLSPRGSRPSNGPTDPRRS	NLG 120	
QY	121	KVDTLTCGFADLMGYTPVVGAPILGGVAAALAHGVRAIEDGINVATGNLPGCCS	SIFLLA 180	
DB	121	KVDTLTCGFADLMGYTPVVGAPILGGVAAALAHGVRAIEDGINVATGNLPGCCS	SIFLLA 180	
QY	181	LLSCLATPASA	191	

```

Db      181  LLSCLTVPASA 191      ||||| ||||
RESULT 13
US-10-365-620-56
; Sequence 56, Application US/10365620
; Publication No. US20040001853A1
; GENERAL INFORMATION:
; APPLICANT: George. Rajan
; APPLICANT: Tyrell, Lorne
; APPLICANT: No. US20040001853A1jaim, Antoine
; TITLE OF INVENTION: Chimeric Antigens for Eliciting An Immune Response
; FILE REFERENCE: 656.0016
; CURRENT APPLICATION NUMBER: US/10/365,620
; PRIOR FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US60/423,578
; PRIOR FILING DATE: 2003-11-05
; PRIOR APPLICATION NUMBER: 60/390,564
; PRIOR FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 473
; TYPE: PRT
; ORGANISM: ORF of HCV Core-TBD protein
US-10-365-620-56

```

Query Match	93.4%	Score 976:	DB 15;	Length 473;
Best Local Similarity	92.7%;	Pred. No.	2.6e-77;	
Matches 177;	Conservative 5;	Mismatches 9;	Indels 0;	Gaps 0
Qy	1	MSTLPKQRKT	KRNTNRPRDVKVPPGGGQIVGGYLLPRRGPRLGVRA	TRKTSERSQPRG 60
Db	31	MSINPKQRKT	KRNTNRPRQDVKFPFGGQIVGGYLLPRGPRLGVA	TRKTSERSQPRG 90
Qy	61	RROPIPKAROPQHWAQGYPMPLYNCEGCWAGWLLSPRGSPHGWPNDRPRS	RNLG 120	
		:   :   :	:   :   :	
Db	91	RRQPIPKARRPEGRTWAPGYPMPLYNCEGCWAGWLLSPRGSRPSWGPTDPRRS	RNLG 150	
Qy	121	KVIDTLTCGFADLMGYIPVVYGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA	180	
Db	151	KVIDTLTCGFADLMGYIPLVGAPLGGGAARALAHGVRVLEDGVNVATGNLPGCSFSIFLLA	210	
Qy	181	LLSCLTTPASA	191	
Db	211	LLSCLTVPASA	221	

RESULT 14  
US-10-912-969-58  
; Sequence 58, Application US/10912969  
; Publication No. US20050013828A1  
; GENERAL INFORMATION:  
; APPLICANT: Virex Research, Inc.  
; APPLICANT: George, Rajan  
; APPLICANT: Tyrrell, Lorne  
; APPLICANT: Noujaim, Antoine  
; APPLICANT: Wang, Dakun  
; APPLICANT: Ma, Allan  
; TITLE OF INVENTION: Chimeric Antigens for Eliciting An Immune Response  
; FILE REFERENCE: 17506-007001  
; CURRENT APPLICATION NUMBER: US/10/912,969  
; CURRENT FILING DATE: 2004-08-05  
; PRIOR APPLICATION NUMBER: US 60/390,564  
; PRIOR FILING DATE: 2002-06-20  
; PRIOR APPLICATION NUMBER: US 60/423,578  
; PRIOR FILING DATE: 2002-11-05  
; PRIOR APPLICATION NUMBER: US 10/365,620  
; PRIOR FILING DATE: 2003-02-13  
; PRIOR APPLICATION NUMBER: PCT/IB04/00373  
; PRIOR FILING DATE: 2004-02-14  
; NUMBER OF SEQ ID NOS: 79



; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 58  
; LENGTH: 473  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
US-10-912-969-58

Query Match 93.4%; Score 976; DB 17; Length 473;  
Best Local Similarity 92.7%; Pred. No. 2.6e-77;  
Matches 177; Conservative 5; Mismatches 9; Indels 0; Gaps 0;  
QY 1 MSTLPKPKQKTKRNTNRRPTDVKFGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
Db |||||  
QY 31 MSTNPKPKQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 90  
Db |||||  
QY 61 RROPIPKARQPGRIHWAQGYWPWLYGNEGCCWAGWLLSPGSRPHWGPNDRRNRNLG 120  
Db |||||  
QY 91 RROPIPKARRPEGRTWAOQGYWPWLYGNEGCCWAGWLLSPGSRPSWGPTDPRRNRNLG 150  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
Db |||||  
QY 151 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 210  
QY 181 LLSCLTTPASA 191  
Db |||||  
QY 211 LLSCLTVPASA 221

RESULT 15  
US-10-913-171-39  
; Sequence 39, Application US/10913171  
; Publication No. US20050031628A1  
; GENERAL INFORMATION:  
; APPLICANT: Virexx Research, Inc.  
; APPLICANT: George, Rajan  
; APPLICANT: Tyrrell, Lorne  
; APPLICANT: Noujaim, Antoine  
; APPLICANT: Wang, Dakun  
; APPLICANT: Ma, Allan  
; TITLE OF INVENTION: CHIMERIC ANTIGENS FOR BREAKING HOST TOLERANCE TO FOREIGN ANTIGENS  
; FILE REFERENCE: 17506-006001  
; CURRENT APPLICATION NUMBER: US/10/913,171  
; CURRENT FILING DATE: 2004-08-05  
; PRIOR APPLICATION NUMBER: US 60/493,449  
; PRIOR FILING DATE: 2004-08-08  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 39  
; LENGTH: 473  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
US-10-913-171-39

Query Match 93.4%; Score 976; DB 17; Length 473;  
Best Local Similarity 92.7%; Pred. No. 2.6e-77;  
Matches 177; Conservative 5; Mismatches 9; Indels 0; Gaps 0;  
QY 1 MSTLPKPKQKTKRNTNRRPTDVKFGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
Db |||||  
QY 31 MSTNPKPKQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 90  
Db |||||  
QY 61 RROPIPKARQPGRIHWAQGYWPWLYGNEGCCWAGWLLSPGSRPHWGPNDRRNRNLG 120  
Db |||||  
QY 91 RROPIPKARRPEGRTWAOQGYWPWLYGNEGCCWAGWLLSPGSRPSWGPTDPRRNRNLG 150  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
Db |||||  
QY 151 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 210

QY 181 LLSCLTTPASA 191  
Db |||||  
QY 211 LLSCLTVPASA 221  
Search completed: July 15, 2005, 10:20:47  
Job time : 166 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2005, 09:53:38 ; Search time 16 Seconds  
(without alignments)  
1148.588 Million cell updates/sec

Title: US-09-084-691B-206  
Perfect score: 1045  
Sequence: 1 MSTLPKPKQKTKRNTNRRPT.....CSFSIFLLALLSCLTTPASA 191  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	976	93.5	3011	1 S40770	genome polyprotein
2	976	93.4	3011	1 GNWVCH	genome polyprotein
3	971	92.9	441	2 S12707	genome polyprotein
4	969	92.7	787	2 S18031	genome polyprotein
5	969	92.7	787	2 PN0677	hypothetical prote
6	968	92.6	3011	1 GNWVC3	genome polyprotein
7	965	92.3	3010	1 GNWVCJ	genome polyprotein
8	965	92.3	3010	1 GNWVC	genome polyprotein
9	962	92.1	550	2 JH0711	genome polyprotein
10	961	92.0	640	2 JQ1584	genome polyprotein
11	959	91.8	189	2 S32740	polyprotein - hepa
12	958	91.7	782	2 S19876	genome polyprotein
13	957	91.6	513	2 PC1284	genome polyprotein
14	957	91.6	782	2 S19875	genome polyprotein
15	957	91.6	876	2 PC2219	polypeptide - hepa
16	956	91.5	3010	1 GNWVTW	genome polyprotein
17	956	91.5	3010	1 S18030	genome polyprotein
18	952	91.1	520	2 JQ1925	polyprotein - hepa
19	951	91.0	411	2 PC2060	genome polyprotein
20	948	90.7	3010	1 A45573	genome polyprotein
21	947	90.6	369	2 S21471	genome polyprotein
22	947	90.6	523	2 JQ1926	polyprotein - hepa
23	947	90.6	782	2 S18032	genome polyprotein
24	946	90.5	492	2 S41288	genome polyprotein
25	944	90.3	411	2 PC2061	genome polyprotein
26	941	90.0	874	2 JQ0883	genome polyprotein
27	940	90.0	513	2 A44150	structural protein
28	939	89.9	874	2 JQ0881	genome polyprotein
29	939	89.9	3033	1 JQ1303	genome polyprotein

RESULT 1

S40770 genome polyprotein - hepatitis C virus  
N:Contains: capsid protein C; envelope protein M; hepatitis virus (nonstructural protein NS4a); nonstructural protein NS4b; nonstructural protein NS5  
C:Species: hepatitis C virus  
C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
C:Accession: S40770; PC1285  
R:Okamoto, H.  
Submitted to the EMBL Data Library, March 1992  
A:Reference number: S40770  
A:Accession: S40770  
A:Molecule type: genomic RNA  
A:Residues: 1-3011 <OK>  
A:Cross-references: UNIPROT:Q03463; EMBL:D10749; NID:g221586; PIDN:BAA01582.1; PID:g221512  
R:Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tauda, Jpn. J. Exp. Med. 60, 167-177, 1990  
A:Title: The 5'-terminal sequence of the hepatitis C virus genome.  
A:Reference number: PC1284; MUID:91013116; PMID:2170712  
A:Accession: PC1285  
A:Molecule type: genomic RNA  
A:Residues: 1-513 <OK2>  
A:Cross-references: GB:D00831; NID:g221511; PIDN:BAA00705.1; PID:g221512  
A:Experimental source: isolate HC-J1  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin  
F:2-115/Product: capsid protein C #status predicted <CPC>  
F:116-191/Product: envelope protein M #status predicted <EPM>  
F:192-389/Product: major envelope protein E #status predicted <MEE>  
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>  
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>  
F:1007-1615/Product: hepatitis virus #status predicted <NS3>  
F:1230-1237/Region: nucleotide-binding motif A (P-loop)  
F:1312-1317/Region: nucleotide-binding motif B  
F:1316-1319/Region: DEXH motif  
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>  
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>  
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match

Best Local Similarity 93.5%; Score 977; DB 1; Length 3011;  
Matches 176; Conservative 92.1%; Pred. No. 2.9e-74;  
Indels 0; Gaps 0;

QY	1	MSTLPKPKQKTKRNTNRRPTDKFPGGQIVGGVYLLPRRGRPRLCVTRATKTSERSQPRG	60
DB	1	MSTLPKPKQKTKRNTNRRPTDKFPGGQIVGGVYLLPRRGRPRLCVTRATKTSERSQPRG	60
QY	61	RRQPTPKARQPGQRHWAQPGYPWPLYNEGCGWAGLLSPRGSRPHMGPNDRPRRSRLNG	120
DB	61	RRQPTPKARQPGQRHWAQPGYPWPLYNEGCGWAGLLSPRGSRPHMGPNDRPRRSRLNG	120
QY	121	KVIDLTCTGCFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGTINATGNLPGCSFIFLLA	180



QY 181 LLSCLTTPASA 191  
|||||  
Db 181 LLSCLTIPASA 191

## RESULT 5

PN0677  
hypothetical protein 787 - hepatitis C virus (fragment)  
C:Species: hepatitis C virus  
C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 09-Jul-2004  
C:Accession: PN0677  
R:Cho, S.H.; Yoon, J.I.; Chang, J.E.; Ahn, B.M.; Lee, C.H.; Lee, Y.I.  
Biochem. Biophys. Res. Commun. 196, 780-788, 1993  
A:Title: Genetic typing of hepatitis C viruses from Korean patients: Implications of genotypes and their relationship to the hepatitis C virus genome  
A:Reference number: PN0677; MUID:94059104; PMID:8240354  
A:Accession: PN0677  
A:Molecule type: mRNA  
A:Residues: 1-787 <CHO>  
A:Cross-references: UNIPROT:Q08244; GB:L20498; NID:g1381031; PIDN:AAB02608.1; PID:g1381031  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: glycoprotein; nonstructural protein  
F:196,209,234,250,305,325,421,427,452,536,544,560,580,627,649/Binding site: carbohydrate binding site  
Query Match 92.7%; Score 969; DB 2; Length 787;  
Best Local Similarity 92.7%; Pred. No. 3.8e-74;  
Matches 177; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 MSTLPKQPKQKRNTRPTDVKFGGGOIVGGVLLPRGRLGVRATRTKTSERSQPRG 60  
|||||  
Db 1 MSTNPKPQKRNTRRPQDVKFGGGOIVGGVLLPRGRLGVRATRTKTSERSQPRG 60  
|||||  
QY 61 RQPIPKARQPKQKRNTRPTDVKFGGGOIVGGVLLPRGRLGVRATRTKTSERSQPRG 120  
|||||  
Db 61 RQPIPKARQPKQKRNTRPTDVKFGGGOIVGGVLLPRGRLGVRATRTKTSERSQPRG 120  
|||||  
QY 121 KVIDTLTCGFADLMGYIPVWGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180  
|||||  
Db 121 KVIDTLTCGFADLMGYIPVWGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180  
|||||  
QY 181 LLSCLTTPASA 191  
|||||  
Db 181 LLSCLTIPASA 191

## RESULT 6

GNWVC3  
Genome polyprotein - hepatitis C virus (strain HCV-1)  
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (strain HCV-1) protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C:Species: hepatitis C virus  
C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 09-Jul-2004  
C:Accession: A39166; PQ0403; FQ0404  
R:Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Coi, Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991  
A:Title: Genetic organization and diversity of the hepatitis C virus.  
A:Reference number: A39166; MUID:91172826; PMID:1848704  
A:Accession: A39166  
A:Molecule type: mRNA  
A:Residues: 1-3011 <CHO>  
A:Cross-references: UNIPROT:P26664; GB:M62321; NID:g329873; PIDN:AAA45676.1; PID:g329874  
R:Chan, S.W.; McMahon, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.I. J. Gen. Virol. 73, 1131-1141, 1992  
A:Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to other hepatitis C virus types  
A:Reference number: PQ0403; MUID:92268871; PMID:1316939  
A:Accession: PQ0403  
A:Molecule type: genomic RNA  
A:Residues: 1577-1633 <CHA>  
A:Cross-references: DDBJ:D10128  
A:Experimental source: isolates E-b16  
A:Accession: PQ0404  
A:Status: preliminary  
A:Molecule type: genomic RNA

A:Residues: 1577-1633 &lt;CH2&gt;

A:Experimental source: isolates E-b17

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural

F:1-115/Product: capsid protein C #status predicted &lt;CPC&gt;

F:116-191/Product: envelope protein M #status predicted &lt;EPM&gt;

F:192-389/Product: major envelope protein E #status predicted &lt;MEB&gt;

F:390-729/Product: nonstructural protein NS1 #status predicted &lt;NS1&gt;

F:730-1006/Product: nonstructural protein NS2 #status predicted &lt;NS2&gt;

F:1007-1615/Product: hepatitis C virus #status predicted &lt;NS3&gt;

F:1230-1237/Region: nucleotide-binding motif A (P-loop)

F:1312-1317/Region: nucleotide-binding motif B

F:1318-1319/Region: DEXH motif

F:1616-1862/Product: nonstructural protein NS4a #status predicted &lt;N4A&gt;

F:1863-2013/Product: nonstructural protein NS4b #status predicted &lt;N4B&gt;

F:2014-3011/Product: nonstructural protein NS5 #status predicted &lt;NS5&gt;

F:196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2077,2078

Query Match 92.6%; Score 968; DB 1; Length 3011;

Best Local Similarity 91.6%; Pred. No. 1.6e-73;

Matches 175; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 MSTLPKQPKQKRNTRPTDVKFGGGOIVGGVLLPRGRLGVRATRTKTSERSQPRG 60

Db 1 MSTNPKPQKRNTRRPQDVKFGGGOIVGGVLLPRGRLGVRATRTKTSERSQPRG 60

QY 61 RQPIPKARQPKQKRNTRPTDVKFGGGOIVGGVLLPRGRLGVRATRTKTSERSQPRG 120

Db 61 RQPIPKARQPKQKRNTRPTDVKFGGGOIVGGVLLPRGRLGVRATRTKTSERSQPRG 120

QY 121 KVIDTLTCGFADLMGYIPVWGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180

Db 121 KVIDTLTCGFADLMGYIPVWGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191

Db 181 LLSCLTIPASA 191

## RESULT 7

GNWVC3  
Genome polyprotein - hepatitis C virus (strain J)  
N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C:Species: hepatitis C virus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
C:Accession: A39253; PS0086  
R:Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimotohno, R. Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990  
A:Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients  
A:Reference number: A39253; MUID:91088550; PMID:2175903  
A:Accession: A39253  
A:Molecule type: genomic RNA  
A:Residues: 1-3010 <KAT>  
A:Cross-references: UNIPROT:P26662; GB:D90208; NID:g221610; PIDN:BA14233.1; PID:g221611  
R:Kato, N.; Ohkoshi, S.; Shimotohno, K. Proc. Jpn. Acad. 65B, 219-223, 1989  
A:Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence variability  
A:Reference number: PS0085  
A:Accession: PS0086  
A:Molecule type: genomic RNA  
A:Residues: 2650-2707 <KA2>  
A:Experimental source: Japanese isolate  
C:Comment: The cleavage sites of this polyprotein have not been determined.  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine protease  
F:2-115/Product: capsid protein C #status predicted <CPC>  
F:116-191/Product: envelope protein M #status predicted <EPM>  
F:192-389/Product: major envelope protein E #status predicted <MEB>  
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>  
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>  
F:1007-1615/Product: hepatitis C virus #status predicted <NS3>  
F:1230-1237/Region: nucleotide-binding motif A (P-loop)



F:390-640/Product: envelope protein E2 and nonstructural protein NS1 #status predicted <F:196,209,234,305,417,430,448,476,540,556,576,623/Binding site: carbohydrate (Asn) (cova

Query Match 92.0%; Score 961; DB 2; Length 640;  
Best Local Similarity 91.1%; Pred. No. 1.5e-73;  
Matches 174; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 MSTLPKPKQRTKNTNRRPTDVKFPGGQIVGGVYLLPRRGPRGLGVRAATKTSERSQPRG 60  
DB 1 MSTNPKPQRTKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRGLGVRAATKTSERSQPRG 60  
QY 61 RROPIPKAROPQGRHWAQPGYPWPYLYGNEGCWAGWLLSPRGRPHWGNDPDRSRNLG 120  
DB 61 RROPIPKARPRPGRTWAQPGYPWPYLYGNEGCWAGWLLSPRGRSPSWGPTDPPRSRNLG 120  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180  
QY 181 LLSCLTTPASA 191  
DB 181 LLSCLTTPASA 191

RESULT 11  
S32740  
polyprotein - hepatitis C virus (isolate Russian) (fragment)  
N:Contains: capsid protein C; envelope protein M  
C:Species: hepatitis C virus  
C:Date: 19-Mar-1997 #sequence\_revision 05-Dec-1998 #text\_change 09-Jul-2004  
C:Accession: S32740  
R:Vassilev, V.B.; Viazov, S.O.; Kotova, E.Y.; Nosikov, V.V.  
submitted to the EMBL Data Library, April 1993  
A:Description: Evidence of new HCV variant of European isolate in Russia.  
A:Reference number: S32740

A:Accession: S32740  
A:Molecule type: genomic RNA  
A:Residues: 1-189 <VAS>  
A:Cross-references: UNIPROT:Q68873; EMBL:X71407  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: capsid protein; envelope protein; polyprotein  
F:1-115/Product: capsid protein C #status predicted <CPC>  
F:116-189/Product: envelope protein M #status predicted <EPM>

Query Match 91.8%; Score 959; DB 2; Length 189;  
Best Local Similarity 92.1%; Pred. No. 6.9e-74;  
Matches 174; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 MSTLPKPKQRTKNTNRRPTDVKFPGGQIVGGVYLLPRRGPRGLGVRAATKTSERSQPRG 60  
DB 1 MSTNPKPQRTKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRGLGVRAATKTSERSQPRG 60  
QY 61 RROPIPKAROPQGRHWAQPGYPWPYLYGNEGCWAGWLLSPRGRPHWGNDPDRSRNLG 120  
DB 61 RROPIPKARPRPGRTWAQPGYPWPYLYGNEGCWAGWLLSPRGRSPSWGPTDPPRSRNLG 120  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180  
QY 181 LLSCLTTPA 189  
DB 181 LLSCLTTPA 189

RESULT 12  
S19876  
genome polyprotein - hepatitis C virus (isolate JK5) (fragment)  
N:Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein  
C:Species: hepatitis C virus  
A:Variety: isolate JK5  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
C:Accession: S19876

R:Honda, M.; Kaneko, S.; Maeshi, U.; Kobayashi, K.; Murakami, S.  
submitted to the EMBL Data Library, September 1991

A:Description: Sequence analysis of putative structural regions of Hepatitis C Virus iso  
A:Reference number: S18029  
A:Accession: S19876  
A:Molecule type: genomic RNA  
A:Residues: 1-782 <HON>

A:Cross-references: UNIPROT:Q68953; EMBL:X61595; NID:g59486; PIDN:CAA43792.1; PID:g59487  
A:Experimental source: isolate JK5  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural  
F:1-391/Product: core protein #status predicted <M11>  
F:192-383/Product: envelope protein 1 #status predicted <MAT2>  
F:384-733/Product: NS1/E2 protein #status predicted <MAT3>  
F:734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Query Match 91.7%; Score 958; DB 2; Length 782;  
Best Local Similarity 91.6%; Pred. No. 3.2e-73;  
Matches 175; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 MSTLPKPKQRTKNTNRRPTDVKFPGGQIVGGVYLLPRRGPRGLGVRAATKTSERSQPRG 60  
DB 1 MSTNPKPQRTKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRGLGVRAATKTSERSQPRG 60  
QY 61 RROPIPKAROPQGRHWAQPGYPWPYLYGNEGCWAGWLLSPRGRPHWGNDPDRSRNLG 120  
DB 61 RROPIPKARPEGRAWAQPGYPWPYLYGNEGCWAGWLLSPRGRSPSWGPTDPPRSRNLG 120  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180  
QY 181 LLSCLTTPASA 191  
DB 181 LLSCLTTPASA 191

RESULT 13  
PC1284

genome polyprotein - hepatitis C virus (isolate HC-J4) (fragment)  
C:Species: hepatitis C virus  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C:Accession: PC1284  
R:Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Jpn. J. Exp. Med. 60, 167-177, 1990  
A:Title: The 5'-terminal sequence of the hepatitis C virus genome.  
A:Reference number: PC1284; MUID:91013116; PMID:2170712

A:Accession: PC1284  
A:Molecule type: genomic RNA  
A:Residues: 1-513 <OKA>  
A:Cross-references: UNIPROT:Q81221; GB:D00832; NID:g221513; PIDN:BAA00706.1; PID:g221514  
C:Superfamily: hepatitis C virus genome polyprotein

Query Match 91.6%; Score 957; DB 2; Length 513;  
Best Local Similarity 91.6%; Pred. No. 2.6e-73;  
Matches 175; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 MSTLPKPKQRTKNTNRRPTDVKFPGGQIVGGVYLLPRRGPRGLGVRAATKTSERSQPRG 60  
DB 1 MSTNPKPQRTKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRGLGVRAATKTSERSQPRG 60  
QY 61 RROPIPKAROPQGRHWAQPGYPWPYLYGNEGCWAGWLLSPRGRPHWGNDPDRSRNLG 120  
DB 61 RROPIPKARPEGRAWAQPGYPWPYLYGNEGCWAGWLLSPRGRSPSWGPTDPPRSRNLG 120  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180  
QY 181 LLSCLTTPASA 191  
DB 181 LLSCLTTPASA 191

RESULT 14  
S19875  
Genome polyprotein - hepatitis C virus (isolate JK3) (fragment)  
N:Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein  
C:Species: hepatitis C virus  
A:Variety: isolate JK3  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
C:Accession: S19875  
R:Honda, M.; Kaneo, S.; Masashi, U.; Kobayashi, K.; Murakami, S.  
submitted to the EMBL Data Library, September 1991  
A:Description: Sequence analysis of putative structural regions of Hepatitis C Virus iso  
A:Reference number: S18029  
A:Accession: S19875  
A:Molecule type: genomic RNA  
A:Residues: 1-782 <HONS>  
A:Cross-references: UNIPROT:Q68951; EMBL:X61592; NID:G59482; PIDN:CAA43789.1; PID:G59483  
A:Experimental source: isolate JK3  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural  
F:1-191/Product: core protein #status predicted <MAT1>  
F:192-383/Product: envelope protein 1 #status predicted <MAT2>  
F:384-733/Product: NS1/E2 protein #status predicted <MAT3>  
F:734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>  
  
Query Match 91.6%; Score 957; DB 2; Length 782;  
Best Local Similarity 91.6%; Pred. No. 3.9e-73;  
Matches 175; Conservative 5; Mismatches 11; Indels 0; Gaps 0;  
  
Qy 1 MSTLPKPQKTKRNTNRRPTDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
Db 1 MSTNPKPQKTKRNTNRRPTDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
  
Qy 61 RRQIPKARQPGGRHWAQPGYPWPPLYNCGWAGWLLSPRGSRPHWGPNDPRRSRLNG 120  
Db 61 RRQIPKARQPGGRHWAQPGYPWPPLYNCGWAGWLLSPRGSRPHWGPNDPRRSRLNG 120  
  
Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180  
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180  
  
Qy 181 LLSCLTTPASA 191  
Db 181 LLSCLTTPASA 191

RESULT 15  
PC2219  
polyptide - hepatitis C virus (type 5a) (fragments)  
N:Contains: core protein; E1 (carboxyl end); E2/NS1 (amino end); NS3 protein; NS4A prote  
C:Species: hepatitis C virus  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: PC2219  
R:Stuyver, L.; Arnhem, W.V.; Wyseur, A.; Maertens, G.  
Biochem. Biophys. Res. Commun. 202, 1308-1314, 1994  
A:Title: Cloning and phylogenetic analysis of the core, E2, and NS3/NS4 regions of the H  
A:Reference number: PC2219; M0ID:94338342; PMID:7520237  
A:Accession: PC2219  
A:Molecule type: mRNA  
A:Residues: 1-876 <STU>  
A:Cross-references: UNIPROT:Q81242; GB:L29577; GB:L29578; GB:L29579  
A:Experimental source: serum  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: glycoprotein  
F:1-191/Product: core #status predicted <COE>  
F:192-247/Product: E1 (carboxyl end) #status predicted <ERE>  
F:248-411/Product: E2/NS1 (amino end) #status predicted <ENR>  
F:248-338/Region: E2  
F:339-411/Region: NS1 (amino end)  
F:412-783/Product: NS3 #status predicted <NSR>  
F:784-837/Product: NS4A #status predicted <NSA>  
F:838-876/Product: NS4B #status predicted <NSB>

F:281,287,294,312,340/Binding site: carbohydrate (Asn) (covalent) #status predicted  
  
Query Match 91.6%; Score 957; DB 2; Length 876;  
Best Local Similarity 90.1%; Pred. No. 4.4e-73;  
Matches 172; Conservative 8; Mismatches 11; Indels 0; Gaps 0;  
  
Qy 1 MSTLPKPQKTKRNTNRRPTDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
Db 1 MSTNPKPQKTKRNTNRRPTDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
  
Qy 61 RRQIPKARQPGGRHWAQPGYPWPPLYNCGWAGWLLSPRGSRPHWGPNDPRRSRLNG 120  
Db 61 RRQIPKARQPGGRHWAQPGYPWPPLYNCGWAGWLLSPRGSRPHWGPNDPRRSRLNG 120  
  
Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180  
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINVATGNLPGCSFSIFLLA 180  
  
Qy 181 LLSCLTTPASA 191  
Db 181 LLSCLTTPASA 191

Search completed: July 15, 2005, 09:59:24  
Job time : 17 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2005, 09:49:58 ; Search time 177 Seconds  
(without alignments)  
552.583 Million cell updates/sec

Title: US-09-084-691B-206  
Perfect score: 1045  
Sequence: 1 MSTLPKPKQKTKRNTNRPT.....CSFSIFLLALLSCLTTPASA 191

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1045	100.0	191	2	Q68115 hepatitis c
2	1035	99.0	191	2	Q68411 hepatitis c
3	1025	98.1	414	2	P89956 hepatitis c
4	1024	98.0	414	2	P89958 hepatitis c
5	1024	98.0	414	2	Q81329 hepatitis c
6	1024	98.0	3016	2	O92531 hepatitis c
7	1023	97.9	3018	2	O39927 hepatitis c
8	1020	97.6	414	2	P89963 hepatitis c
9	1016	97.2	414	2	P89957 hepatitis c
10	1016	97.2	414	2	P89960 hepatitis c
11	1014	97.0	414	2	P89959 hepatitis c
12	1011	96.7	414	2	Q81270 hepatitis c
13	1009	96.6	319	2	Q68703 hepatitis c
14	1009	96.6	414	2	P89962 hepatitis c
15	1007	96.4	191	2	O93057 hepatitis c
16	1007	96.4	414	2	Q81269 hepatitis c
17	1003	96.0	414	2	Q81268 hepatitis c
18	1003	96.0	414	2	Q81314 hepatitis c
19	1001	95.8	191	2	O70635 hepatitis c
20	1000	95.7	414	2	Q81271 hepatitis c
21	999	95.6	414	2	P89955 hepatitis c
22	997	95.4	414	2	P89954 hepatitis c
23	997	95.4	415	2	P89961 hepatitis c
24	997	95.4	415	2	P89964 hepatitis c
25	997	95.4	415	2	P89965 hepatitis c
26	997	95.4	3015	2	O92532 hepatitis c
27	996	95.3	191	2	Q68153 hepatitis c
28	995	95.2	415	2	Q81272 hepatitis c
29	995	95.2	415	2	Q81315 hepatitis c
30	990	94.7	415	2	P87751 hepatitis c
31	990	94.7	3013	2	O92530 hepatitis c

32	988	94.5	414	2	Q81267 hepatitis c
33	988	94.5	514	2	Q6XJ47 hepatitis c
34	988	94.5	3019	2	O92529 hepatitis c
35	987	94.4	191	2	P87841 hepatitis c
36	987	94.4	415	2	Q81274 hepatitis c
37	985	94.3	416	2	O39647 hepatitis c
38	984	94.2	319	2	Q86698 hepatitis c
39	984	94.2	415	2	Q81273 hepatitis c
40	982	94.0	191	2	Q68152 hepatitis c
41	982	94.0	191	2	Q68157 hepatitis c
42	982	94.0	319	2	Q68705 hepatitis c
43	981	93.9	191	2	Q68116 hepatitis c
44	981	93.9	191	2	Q68576 hepatitis c
45	981	93.9	326	2	Q81257 hepatitis c

## ALIGNMENTS

### RESULT 1

Q68115 PRELIMINARY; PRT; 191 AA.  
AC Q68115;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Polypeptide (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP MEDLINE=94336721; PubMed=8058787;  
RA Bukh J., Purcell R.H., Miller R.H.;  
RT "Sequence analysis of the core gene of 14 hepatitis C virus genotypes."  
RL Proc. Natl. Acad. Sci. U.S.A. 91:8239-8243(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92279243; PubMed=1317578;  
RA Bukh J., Purcell R.H., Miller R.H.;  
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus."  
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93376778; PubMed=8396266;  
RA Bukh J., Purcell R.H., Miller R.H.;  
RT "At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative E1 gene of isolates collected worldwide."  
RL Proc. Natl. Acad. Sci. U.S.A. 90:8234-8238(1993).  
DR EMBL; U10198; AAA21037.1;  
DR HSP; Q8JYS1; 1CWK.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR002522; HCV capsid.  
DR InterPro; IPR002521; HCV core.  
DR Pfam; PF01543; HCV capsid; 1.  
DR Pfam; PF01542; HCV\_core; 1.  
KW Polyprotein.  
FT CHAIN 1 >191 core protein.  
FT NON TER 191  
SQ SEQUENCE 191 AA; 20725 MW; 7A7994DF2F0909C6 CRC64;

Query Match 100.0%; Score 1045; DB 2; Length 191;  
Best Local Similarity 100.0%; Pred. No. 3e-80;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSTLPKPKQKTKRNTNRPTDKVPGGQIVGGVYLLPRRGPRGVTRATKTSERSQPRG 60
Db	1	MSTLPKPKQKTKRNTNRPTDKVPGGQIVGGVYLLPRRGPRGVTRATKTSERSQPRG 60
QY	61	RRQIPKARQFQGRHWAQPGYPWPVLYGNEGCGWLLSPRGSRPHWGFNDPRRSRLNG 120

```

|||||
Db 61 RROQIPKARQPGQRHWAQPGYVPWLYGNEGCGWAGWLLSPRSGRPHWGPNDRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 2
Q68411 PRELIMINARY; PRT; 191 AA.
AC Q68411;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Core protein (Fragment)
OS Hepatitis C virus type 6.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=42182;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97052554; PubMed=8897188;
RA Bernier L., Willes B., Delage G., Murphy D.G.;
RT "Identification of numerous hepatitis C virus genotypes in Montreal,
Canada.";
RL J. Clin. Microbiol. 34:2815-2818(1996).
DR EMBL; U33435; AAB40038.1; -.
DR HSSP; Q8JYSL; ICWX.
GO GO:0019028; C:viral capsid; IEA.
GO GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002521; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
FT NON_TER 191
SQ SEQUENCE 191 AA; 20789 MW; F292AB64B56DE30A CRC64;

Query Match 99.0%; Score 1035; DB 2; Length 191;
Best Local Similarity 99.0%; Pred. No. 2.1e-79;
Matches 189; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTLPKQKTKRNTNRRPTDVKFGGGQIVGGVYLLPRRGLGVRAIETGNYATGNLPGCSFSIFLLA 60
Db 1 MSTLPKQKTKRNTNRRPMDVKFGGGQIVGGVYLLPRRGLGVRAIETGNYATGNLPGCSFSIFLLA 60
QY 61 RROQIPKARQPGQRHWAQPGYVPWLYGNEGCGWAGWLLSPRSGRPHWGPNDRSRNLG 120
Db 61 RROQIPKARQPGQRHWAQPGYVPWLYGNEGCGWAGWLLSPRSGRPHWGPNDRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 3
P89956 PRELIMINARY; PRT; 414 AA.
ID P89956
AC P89956;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

```

```

OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95062197; PubMed=7972001;
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
Iizuka H., Mishiro S., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Vietnam are classifiable into the
seventh, eighth, and ninth major genetic groups.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Okamoto H.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D88469; BAAL3618.1; -.
DR PIR; PQ0804; PQ0804.
DR HSSP; Q8JYSL; ICWX.
GO GO:0016021; C:integral to membrane; IEA.
GO GO:0019028; C:viral capsid; IEA.
GO GO:0019031; C:viral envelope; IEA.
GO GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; Transmembrane.
FT NON_TER 414
SQ SEQUENCE 414 AA; 44528 MW; 7A3BE1710311C017 CRC64;

Query Match 98.1%; Score 1025; DB 2; Length 414;
Best Local Similarity 97.9%; Pred. No. 3.2e-78;
Matches 187; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSTLPKQKTKRNTNRRPTDVKFGGGQIVGGVYLLPRRGLGVRAIETGNYATGNLPGCSFSIFLLA 60
Db 1 MSTLPKQKTKRNTNRRPMDVKFGGGQIVGGVYLLPRRGLGVRAIETGNYATGNLPGCSFSIFLLA 60
QY 61 RROQIPKARQPGQRHWAQPGYVPWLYGNEGCGWAGWLLSPRSGRPHWGPNDRSRNLG 120
Db 61 RROQIPKARQPGQRHWAQPGYVPWLYGNEGCGWAGWLLSPRSGRPHWGPNDRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASA 191
Db 181 LLSCLTTPASA 191

RESULT 4
P89958 PRELIMINARY; PRT; 414 AA.
ID P89958
AC P89958;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95062197; PubMed=7972001;
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
Iizuka H., Mishiro S., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Vietnam are classifiable into the

```

RT seventh, eighth, and ninth major genetic groups.";  
RN Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).  
RP SEQUENCE FROM N.A.  
RA Okamoto H.;  
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; D88471; BAAL13620.1; --  
DR HSSP; Q8JYS1; ICWX.  
DR GO:0016021; C:integral to membrane; IEA.  
DR GO:0019028; C:viral capsid; IEA.  
DR GO:0019031; F:structural molecule; IEA.  
DR GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR002522; HCV capsid.  
DR InterPro; IPR002521; HCV\_core.  
DR InterPro; IPR002531; HCV\_env.  
DR Pfam; PF01543; HCV capsid; 1.  
DR Pfam; PF01542; HCV core; 1.  
DR Pfam; PF01539; HCV env; 1.  
DR Pfam; PF01560; HCV NS1; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
Polyprotein; Transmembrane.  
FT NON\_TER 414  
SQ SEQUENCE 414 AA; 44973 MW; F3F3CF154372FAFF CRC64;  
  
Query Match 98.0%; Score 1024; DB 2; Length 414;  
Best Local Similarity 97.4%; Pred. No. 3.9e-78;  
Matches 186; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 MSTLPKPKQKRNTRRPTDVKFGGGQIVGGVLLPRRGLGVRAIRTKTSERSQPRG 60  
DB 1 MSTLPKPKQKRNTRRPMQVKGPGGQIVGGVLLPRRGLGVRAIRTKTSERSQPRG 60  
  
QY 61 RRQIPKARQPGRHWAQGYWPPLYGNEGCGWAGWLLSPRSGRPHWGPNDRPRSRNLG 120  
DB 61 RRQIPKARQPGRHWAQGYWPPLYGNEGCGWAGWLLSPRSGRPHWGPNDRPRSRNLG 120  
  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180  
  
QY 181 LLSCLTTPASA 191  
DB 181 LLSCLTTPASA 191  
  
RESULT 5  
Q81329 PRELIMINARY; PRT; 414 AA.  
AC Q81329;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Core, env and part of E2/NS1 (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=111103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95062197; PubMed=7972001;  
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,  
Iizuka H., Mishihiro S., Miyakawa Y., Mayumi M.;  
RT "Hepatitis C virus variants from Vietnam are classifiable into the  
seventh, eighth, and ninth major genetic groups";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Okamoto H.;  
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; D88468; BAAL13617.1; --  
DR FIR; PQ0804; PQ0804.  
DR HSSP; Q8JYS1; ICWX.

DR GO:0016021; C:integral to membrane; IEA.  
DR GO:0019028; C:viral capsid; IEA.  
DR GO:0019031; C:viral envelope; IEA.  
DR GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR002522; HCV capsid.  
DR InterPro; IPR002521; HCV\_core.  
DR InterPro; IPR002531; HCV\_env.  
DR Pfam; PF01543; HCV capsid; 1.  
DR Pfam; PF01542; HCV core; 1.  
DR Pfam; PF01539; HCV env; 1.  
DR Pfam; PF01560; HCV NS1; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
Polyprotein; Transmembrane.  
FT NON\_TER 414  
SQ SEQUENCE 414 AA; 44937 MW; 65355640863B3DCP CRC64;  
  
Query Match 98.0%; Score 1024; DB 2; Length 414;  
Best Local Similarity 98.4%; Pred. No. 3.9e-78;  
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 MSTLPKPKQKRNTRRPTDVKFGGGQIVGGVLLPRRGLGVRAIRTKTSERSQPRG 60  
DB 1 MSTLPKPKQKRNTRRPMQVKGPGGQIVGGVLLPRRGLGVRAIRTKTSERSQPRG 60  
  
QY 61 RRQIPKARQPGRHWAQGYWPPLYGNEGCGWAGWLLSPRSGRPHWGPNDRPRSRNLG 120  
DB 61 RRQIPKARQPGRHWAQGYWPPLYGNEGCGWAGWLLSPRSGRPHWGPNDRPRSRNLG 120  
  
QY 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180  
  
QY 181 LLSCLTTPASA 191  
DB 181 LLSCLTTPASA 191  
  
RESULT 6  
O92531 PRELIMINARY; PRT; 3016 AA.  
AC O92531;  
DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE Polyprotein.  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=111103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98378034; PubMed=9714232;  
RA Tokita H., Okamoto H., Iizuka H., Kishimoto J., Tsuda F., Miyakawa Y.,  
Mayumi M.;  
RT "The entire nucleotide sequences of three hepatitis C virus isolates  
in genetic groups 7-9 and comparison with those in the other eight  
genetic groups";  
RL J. Gen. Virol. 79:1847-1857(1998).  
DR EMBL; D84264; BAA32666.1; --  
DR FIR; PQ0804; PQ0804.  
DR HSSP; Q8JYS1; ICWX.  
DR MEROPS; S29.001; --  
DR MEROPS; U39.001; --  
DR GO:0016021; C:integral to membrane; IEA.  
DR GO:0019028; C:viral capsid; IEA.  
DR GO:0019031; C:viral envelope; IEA.  
DR GO:0005524; F:ATP binding; IEA.  
DR GO:0008026; F:ATP-dependent helicase activity; IEA.  
DR GO:0003723; F:RNA binding; IEA.  
DR GO:0003723; F:RNA binding; IEA.  
DR GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
DR GO:0008236; F:serine-type peptidase activity; IEA.  
DR GO:0005198; F:structural molecule activity; IEA.

DR GO: 0006508; P:proteolysis and peptidolysis; IEA.  
 DR GO: 0006350; P:transcription; IEA.  
 DR GO: 0019079; P:viral genome replication; IEA.  
 DR GO: 0019087; P:viral transformation; IEA.  
 DR InterPro: IPR000345; CytC\_heme\_BS.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR011545; DEAD/DEAH\_N.  
 DR InterPro: IPR002522; HCV\_capsid.  
 DR InterPro: IPR002521; HCV\_core.  
 DR InterPro: IPR002519; HCV env.  
 DR InterPro: IPR002531; HCV NS1.  
 DR InterPro: IPR000745; HCV NS4a.  
 DR InterPro: IPR001490; HCV NS4b.  
 DR InterPro: IPR002868; HCV NS5a.  
 DR InterPro: IPR002166; HCV RdRP.  
 DR InterPro: IPR004109; Peptidase\_S29.  
 DR InterPro: IPR009003; Pept\_Ser\_Cys.  
 DR InterPro: IPR002518; Pept\_U39\_HCV\_NS2.  
 DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro: IPR007094; RNA\_pol\_PSvir.  
 DR Pfam: PF01543; HCV capsid; 1.  
 DR Pfam: PF01542; HCV core; 1.  
 DR Pfam: PF01539; HCV env; 1.  
 DR Pfam: PF01560; HCV NS1; 1.  
 DR Pfam: PF01538; HCV NS2; 1.  
 DR Pfam: PF02907; HCV NS3; 1.  
 DR Pfam: PF01006; HCV NS4a; 1.  
 DR Pfam: PF01001; HCV NS4b; 1.  
 DR Pfam: PF01506; HCV NS5a; 1.  
 DR Pfam: PF00998; Viral RdRP; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN\_1.  
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 SQ SEQUENCE 3016 AA; 328032 MW; 4B5CF96258BCE3B CRC64;

Query Match 98.0%; Score 1024; DB 2; Length 3016;  
 Best Local Similarity 98.4%; Pred. No. 2.8e-77;  
 Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKNTNRRPTDVKFPGGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60  
 DB 1 MSTLPKPKQKTKNTNRRPMDVKFPGGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60

QY 61 RQQIPKARQPGQRHWAQPGYPWPLYNCGWAGWLLSPRGRPHWGNDPDRSRNLG 120  
 DB 61 RQQIPKARQPGQRHWAQPGYPWPLYNCGWAGWLLSPRGRPHWGNDPDRSRNLG 120

QY 121 KVDTLTGCFADLMGYIPVVGAPLGVAALAHGVRAIEDGYNATGNLPGCSFSIFLLA 180  
 DB 121 KVDTLTGCFADLMGYIPVVGAPLGVAALAHGVRAIEDGYNATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191  
 DB 181 LLSCLTTPASA 191

RESULT 7  
 O39927 ID O39927 PRELIMINARY; PRT; 3018 AA.  
 AC O39927;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Polyprotein.  
 OS Hepatitis C virus type 6a.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus; Hepatitis C virus type 6.  
 OX NCBI\_Taxid=31655;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=euhk2;  
 RX MEDLINE=97320431; PubMed=9177282; DOI=10.1006/bbrc.1997.6627;

RA Adams A., Chamberlain R.W., Taylor L.A., Davidson F., Lin C.K.,  
 RA Simmonds P., Elliot R.M.;  
 RT "Complete coding sequence of hepatitis C virus genotype 6a.";  
 RL Biochem. Biophys. Res. Commun. 234:393-396(1997).  
 DR EMBL; Y12083; CAA72801.1; -.  
 DR HSSP; Q8JYS1; 1CWK.  
 DR MEROPS; S29.001; -.  
 DR MEROPS; U39.001; -.  
 DR GO: 0016021; C: integral to membrane; IEA.  
 DR GO: 0019028; C: viral capsid; IEA.  
 DR GO: 0019031; C: viral envelope; IEA.  
 DR GO: 0005524; F: ATP binding; IEA.  
 DR GO: 0008026; F: ATP-dependent helicase activity; IEA.  
 DR GO: 0003723; F: RNA binding; IEA.  
 DR GO: 0003968; F: RNA-directed RNA polymerase activity; IEA.  
 DR GO: 0008236; F: serine-type peptidase activity; IEA.  
 DR GO: 0005198; F: structural molecule activity; IEA.  
 DR GO: 0006508; P: proteolysis and peptidolysis; IEA.  
 DR GO: 0006350; P: transcription; IEA.  
 DR GO: 0019079; P: viral genome replication; IEA.  
 DR GO: 0019087; P: viral transformation; IEA.  
 DR InterPro: IPR000345; CytC\_heme\_BS.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR011545; DEAD/DEAH\_N.  
 DR InterPro: IPR002522; HCV\_capsid.  
 DR InterPro: IPR002521; HCV\_core.  
 DR InterPro: IPR002519; HCV env.  
 DR InterPro: IPR002531; HCV NS1.  
 DR InterPro: IPR000745; HCV NS4a.  
 DR InterPro: IPR001490; HCV NS4b.  
 DR InterPro: IPR002868; HCV NS5a.  
 DR InterPro: IPR002166; HCV RdRP.  
 DR InterPro: IPR004109; Peptidase\_S29.  
 DR InterPro: IPR009003; Pept\_Ser\_Cys.  
 DR InterPro: IPR002518; Pept\_U39\_HCV\_NS2.  
 DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro: IPR007094; RNA\_pol\_PSvir.  
 DR Pfam: PF01543; HCV capsid; 1.  
 DR Pfam: PF01542; HCV core; 1.  
 DR Pfam: PF01539; HCV env; 1.  
 DR Pfam: PF01560; HCV NS1; 1.  
 DR Pfam: PF01538; HCV NS2; 1.  
 DR Pfam: PF02907; HCV NS3; 1.  
 DR Pfam: PF01006; HCV NS4a; 1.  
 DR Pfam: PF01001; HCV NS4b; 1.  
 DR Pfam: PF01506; HCV NS5a; 1.  
 DR Pfam: PF00998; Viral RdRP; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN\_1.  
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 SQ SEQUENCE 3018 AA; 329017 MW; 6E67FB3CF6A61AE3 CRC64;

Query Match 97.9%; Score 1023; DB 2; Length 3018;  
 Best Local Similarity 97.9%; Pred. No. 3.4e-77;  
 Matches 187; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKNTNRRPTDVKFPGGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60  
 DB 1 MSTLPKPKQKTKNTNRRPMDVKFPGGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60

QY 61 RQQIPKARQPGQRHWAQPGYPWPLYNCGWAGWLLSPRGRPHWGNDPDRSRNLG 120  
 DB 61 RQQIPKARQPGQRHWAQPGYPWPLYNCGWAGWLLSPRGRPHWGNDPDRSRNLG 120

QY 121 KVDTLTGCFADLMGYIPVVGAPLGVAALAHGVRAIEDGYNATGNLPGCSFSIFLLA 180  
 DB 121 KVDTLTGCFADLMGYIPVVGAPLGVAALAHGVRAIEDGYNATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191  
 DB 181 LLSCLTTPASA 191

```
RESULT 8
P89963 ID P89963 PRELIMINARY; PRT; 414 AA.
AC P89963;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95062197; PubMed=7972001;
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
RA Iizuka H., Mishihiro S., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Vietnam are classifiable into the
RT seventh, eighth, and ninth major genetic groups.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Okamoto H.;
RX Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D88476; BAA13625.1; -.
DR HSSP; Q8JYS1; 1CWV.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01539; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 414
SQ SEQUENCE 414 AA; 44804 MW; 397D0DC54BDA20D CRC64;

Query Match 97.6%; Score 1020; DB 2; Length 414;
Best Local Similarity 96.9%; Pred. No. 8.4e-78;
Matches 185; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSTLPKQKTKRNTNRRPTDKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
DB 1 MSTLPKQKTKRNTNRRPTDKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60

QY 61 RQPIPKARQPOGRHWAQPGYWPVLYGNEGCGWAGLLSPRGRPHWGPNDRSRNLG 120
DB 61 RQPIPKARQPOGRHWAQPGYWPVLYGNEGCGWAGLLSPRGRPHWGPNDRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 9
P89957 ID P89957 PRELIMINARY; PRT; 414 AA.
AC P89957;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95062197; PubMed=7972001;
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
RA Iizuka H., Mishihiro S., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Vietnam are classifiable into the
RT seventh, eighth, and ninth major genetic groups.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Okamoto H.;
RX Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D88476; BAA13625.1; -.
DR HSSP; Q8JYS1; 1CWV.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01539; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 414
SQ SEQUENCE 414 AA; 44804 MW; 397D0DC54BDA20D CRC64;

Query Match 97.6%; Score 1020; DB 2; Length 414;
Best Local Similarity 96.9%; Pred. No. 8.4e-78;
Matches 185; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSTLPKQKTKRNTNRRPTDKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
DB 1 MSTLPKQKTKRNTNRRPTDKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60

QY 61 RQPIPKARQPOGRHWAQPGYWPVLYGNEGCGWAGLLSPRGRPHWGPNDRSRNLG 120
DB 61 RQPIPKARQPOGRHWAQPGYWPVLYGNEGCGWAGLLSPRGRPHWGPNDRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191
```

```
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95062197; PubMed=7972001;
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
RA Iizuka H., Mishihiro S., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Vietnam are classifiable into the
RT seventh, eighth, and ninth major genetic groups.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Okamoto H.;
RX Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D88476; BAA13619.1; -.
DR HSSP; Q8JYS1; 1CWV.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 414
SQ SEQUENCE 414 AA; 44850 MW; 1754DDD583F52414 CRC64;

Query Match 97.2%; Score 1016; DB 2; Length 414;
Best Local Similarity 96.9%; Pred. No. 1.8e-77;
Matches 185; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSTLPKQKTKRNTNRRPTDKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
DB 1 MSTLPKQKTKRNTNRRPTDKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60

QY 61 RQPIPKARQPOGRHWAQPGYWPVLYGNEGCGWAGLLSPRGRPHWGPNDRSRNLG 120
DB 61 RQPIPKARQPOGRHWAQPGYWPVLYGNEGCGWAGLLSPRGRPHWGPNDRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191
DB 181 LLSCLTTPASA 191

RESULT 10
P89960 ID P89960 PRELIMINARY; PRT; 414 AA.
AC P89960;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Core, env and part of E2/NS1 (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95062197; PubMed=7972001;
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
RA Iizuka H., Mishihiro S., Miyakawa Y., Mayumi M.;
```

RT "Hepatitis C virus variants from Vietnam are classifiable into the  
 RT seventh, eighth, and ninth major genetic groups.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Okamoto H.;  
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; D88473; BAA13622.1; -.  
 DR F01; PQ0804; PQ0804.  
 DR HSP; O8JYS1; 1CWK.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 RA InterPro; IPR002522; HCV capsid.  
 DR InterPro; IPR002521; HCV\_core.  
 DR InterPro; IPR002531; HCV\_nsl.  
 DR Pfam; PF01543; HCV\_capsid; 1.  
 DR Pfam; PF01539; HCV\_env; 1.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR002522; HCV capsid.  
 DR InterPro; IPR002521; HCV\_core.  
 DR InterPro; IPR002531; HCV\_nsl.  
 DR Pfam; PF01543; HCV\_capsid; 1.  
 DR Pfam; PF01539; HCV\_env; 1.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 FT NON\_TER 414 414  
 SQ SEQUENCE 414 AA; 44630 MW; C7DDC84D0198BBD7 CRC64;  
 Query Match 97.2%; Score 1016; DB 2; Length 414;  
 Best Local Similarity 96.9%; Pred. No. 1.8e-77;  
 Matches 185; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MSTLPKQKTKRNTNRPTDVKPGGQIVGGVLLPRGPRLGVRATRTKTSERSQPRG 60  
 DB 1 MSTLPKQKTKRNTNRPTDVKPGGQIVGGVLLPRGPRLGVRATRTKTSERSQPRG 60  
 QY 61 RQPIPKARQPGQRHWAQGYWPVLYGNEGCGWAGWLLSPRSGRPHWGNDPRRSRLG 120  
 DB 61 RQPIPKARQPGQRHWAQGYWPVLYGNEGCGWAGWLLSPRSGRPHWGNDPRRSRLG 120  
 QY 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
 DB 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
 QY 181 LLSCLTTPASA 191  
 DB 181 LLSCLTTPASA 191  
 RESULT 11  
 P89959 PRELIMINARY; PRT; 414 AA.  
 ID P89959  
 AC P89959  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Core, env and part of E2/NS1 (Fragment).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95062197; PubMed=7972001;  
 RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,  
 RA Iizuka H., Mishiro S., Miyakawa Y., Mayumi M.;  
 RT "Hepatitis C virus variants from Vietnam are classifiable into the  
 RT seventh, eighth, and ninth major genetic groups.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Okamoto H.;  
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; F88472; BAA13621.1; -.  
 Query Match 97.2%; Score 1016; DB 2; Length 414;  
 Best Local Similarity 96.9%; Pred. No. 1.8e-77;  
 Matches 185; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MSTLPKQKTKRNTNRPTDVKPGGQIVGGVLLPRGPRLGVRATRTKTSERSQPRG 60  
 DB 1 MSTLPKQKTKRNTNRPTDVKPGGQIVGGVLLPRGPRLGVRATRTKTSERSQPRG 60  
 QY 61 RQPIPKARQPGQRHWAQGYWPVLYGNEGCGWAGWLLSPRSGRPHWGNDPRRSRLG 120  
 DB 61 RQPIPKARQPGQRHWAQGYWPVLYGNEGCGWAGWLLSPRSGRPHWGNDPRRSRLG 120  
 QY 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
 DB 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
 QY 181 LLSCLTTPASA 191  
 DB 181 LLSCLTTPASA 191  
 RESULT 11  
 P89959 PRELIMINARY; PRT; 414 AA.  
 ID P89959  
 AC P89959  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Core, env and part of E2/NS1 (Fragment).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95062197; PubMed=7972001;  
 RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,  
 RA Iizuka H., Mishiro S., Miyakawa Y., Mayumi M.;  
 RT "Hepatitis C virus variants from Vietnam are classifiable into the  
 RT seventh, eighth, and ninth major genetic groups.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Okamoto H.;  
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; F88472; BAA13621.1; -.  
 Query Match 97.0%; Score 1014; DB 2; Length 414;  
 Best Local Similarity 96.9%; Pred. No. 2.7e-77;  
 Matches 185; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MSTLPKQKTKRNTNRPTDVKPGGQIVGGVLLPRGPRLGVRATRTKTSERSQPRG 60  
 DB 1 MSTLPKQKTKRNTNRPTDVKPGGQIVGGVLLPRGPRLGVRATRTKTSERSQPRG 60  
 QY 61 RQPIPKARQPGQRHWAQGYWPVLYGNEGCGWAGWLLSPRSGRPHWGNDPRRSRLG 120  
 DB 61 RQPIPKARQPGQRHWAQGYWPVLYGNEGCGWAGWLLSPRSGRPHWGNDPRRSRLG 120  
 QY 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
 DB 121 KVIDTLTCGFADLMGYIPVVGAPLGVAALAHGVRAIEDGINYATGNLPGCSFSIFLLA 180  
 QY 181 LLSCLTTPASA 191  
 DB 181 LLSCLTTPASA 191  
 RESULT 12  
 Q81270 PRELIMINARY; PRT; 414 AA.  
 ID Q81270  
 AC Q81270  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Core, env and part of E2/NS1 (Fragment).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96005057; PubMed=7561773;  
 RA Tokita H., Okamoto H., Luengrojjanakul P., Vareesangthip K.,  
 RA Chainuvatti T., Iizuka H., Tsuda F., Miyakawa Y., Mayumi M.;  
 RT "Hepatitis C virus variants from Thailand classifiable into five novel  
 RT genotypes in the sixth (6b), seventh (7c,7d) and ninth (9b, 9c) major  
 RT genetic groups.";  
 RL J. Gen. Virol. 76:2329-2335(1995).  
 DR EMBL; D37845; BAA07091.1; -.  
 DR HSP; O8JYS1; 1CWK.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR002522; HCV capsid.  
 DR InterPro; IPR002521; HCV\_core.  
 DR InterPro; IPR002531; HCV\_nsl.  
 DR Pfam; PF01543; HCV\_capsid; 1.  
 DR Pfam; PF01539; HCV\_env; 1.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 FT NON\_TER 414 414  
 SQ SEQUENCE 414 AA; 44892 MW; 695600C405A2C082 CRC64;

DR Pfam; PF01560; HCV NS1; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT NON\_TER 414 414  
SQ SEQUENCE 414 AA; 44994 MW; 8C1E123FC61D89FC CRC64;

Query Match 96.7%; Score 1011; DB 2; Length 414;  
Best Local Similarity 97.4%; Pred. No. 4.8e-77;  
Matches 186; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRPTDVKPGGGQIVGGVYLLPRRGLRVATRKTSERSQPRG 60  
DB 1 MSTLPKPKQKTKRNTNRPTDVKPGGGQIVGGVYLLPRRGLRVATRKTSERSQPRG 60

QY 61 RQPIPKARQPGQGRHWAQPGYWPPLYGNEGCGWAGWLLSPRGRPHWGPNDPRRSRNILG 120  
DB 61 RQPIPKARQPGQGRHWAQPGYWPPLYGNEGCGWAGWLLSPRGRPHWGPNDPRRSRNILG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191  
DB 181 LLSCLTTPASA 191

RESULT 13

Q68703 PRELIMINARY; PRT; 319 AA.

DT 01-NOV-1996 (TREMELrel. 01, Created)  
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
DE Core protein (Fragment)  
OS Hepatitis C virus type 6a.  
OC Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus; Hepatitis C virus type 6.  
OX NCBI\_TaxID=31655;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96118171; PubMed=8578855; DOI=10.1016/0168-1702(95)00052-R;  
RA Stuyver L., Wyseur A., van Arnhem W., Lunel F., Laurent-Puig P.,  
RA Pawlotsky J.-M., Kleter B., Baesit L., Nkengasong J., van Doorn L.-J.,  
RA Maertens G.;  
RT "Hepatitis C virus genotyping by means of 5'-UR/core line probe assays  
RT and molecular analysis of untypeable samples.";  
RL Virus Res. 38:137-157(1995).  
DR EMBL; L38339; AAC42193.1; -.  
DR HSSP; Q8JY81; 1CWK.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:vital capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR002521; HCV capsid.  
DR InterPro; IPR002519; HCV env.  
DR Pfam; PF01543; HCV capsid; 1.  
DR Pfam; PF01539; HCV env; 1.  
DR Pfam; PF01560; HCV NS1; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
KW Transmembrane.  
FT CHAIN 1 191 core protein.  
FT CHAIN 192 >319 E1 protein.  
FT NON\_TER 319 319  
SQ SEQUENCE 319 AA; 34545 MW; DC1D0EA52ED64A8F CRC64;

Query Match 96.6%; Score 1009; DB 2; Length 319;  
Best Local Similarity 96.3%; Pred. No. 5.4e-77;  
Matches 184; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRPTDVKPGGGQIVGGVYLLPRRGLRVATRKTSERSQPRG 60  
DB 1 MSTLPKPKQKTKRNTNRPTDVKPGGGQIVGGVYLLPRRGLRVATRKTSERSQPRG 60

DB 1 MSTLPKPKQKTKRNTNRPTDVKPGGGQIVGGVYLLPRRGLRVATRKTSERSQPRG 60  
QY 61 RQPIPKARQPGQGRHWAQPGYWPPLYGNEGCGWAGWLLSPRGRPHWGPNDPRRSRNILG 120  
DB 61 RQPIPKARQPGQGRHWAQPGYWPPLYGNEGCGWAGWLLSPRGRPHWGPNDPRRSRNILG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180

QY 181 LLSCLTTPASA 191  
DB 181 LLSCLTTPASA 191

RESULT 14

P89962 PRELIMINARY; PRT; 414 AA.

AC P89962;  
DT 01-MAY-1997 (TREMELrel. 03, Created)  
DT 01-MAY-1997 (TREMELrel. 03, Last sequence update)  
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
DE Core, env and part of E2/NS1 (Fragment).  
OS Hepatitis C virus.  
OC Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=111103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95062197; PubMed=7972001;  
RA Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,  
RA Izuka H., Mishihiro S., Miyakawa Y., Mayumi M.;  
RT "Hepatitis C virus variants from Vietnam are classifiable into the  
RT seventh, eighth, and ninth major genetic groups.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Okamoto H.;  
RA Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; D88475; BAA13624.1; -.  
DR HSSP; Q8JY81; 1CWK.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:vital capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR002522; HCV capsid.  
DR InterPro; IPR002521; HCV core.  
DR InterPro; IPR002519; HCV env.  
DR InterPro; IPR002531; HCV NS1.  
DR Pfam; PF01543; HCV capsid; 1.  
DR Pfam; PF01542; HCV core; 1.  
DR Pfam; PF01539; HCV env; 1.  
DR Pfam; PF01560; HCV NS1; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT NON\_TER 414 414  
SQ SEQUENCE 414 AA; 44751 MW; 4BBA6143BD11E3C6 CRC64;

Query Match 96.6%; Score 1009; DB 2; Length 414;  
Best Local Similarity 96.3%; Pred. No. 7.1e-77;  
Matches 184; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSTLPKPKQKTKRNTNRPTDVKPGGGQIVGGVYLLPRRGLRVATRKTSERSQPRG 60  
DB 1 MSTLPKPKQKTKRNTNRPTDVKPGGGQIVGGVYLLPRRGLRVATRKTSERSQPRG 60

QY 61 RQPIPKARQPGQGRHWAQPGYWPPLYGNEGCGWAGWLLSPRGRPHWGPNDPRRSRNILG 120  
DB 61 RQPIPKARQPGQGRHWAQPGYWPPLYGNEGCGWAGWLLSPRGRPHWGPNDPRRSRNILG 120

QY 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPVVGAPLGGVAAALAHGVRAIEDGINATGNLPGCSFSIFLLA 180





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2005, 09:59:09 ; Search time 160 Seconds  
(without alignments)  
461.695 Million cell updates/sec

Title: US-09-084-691B-206  
Perfect score: 191  
Sequence: 1 MDTLPKPKRKNTRRRPT.....CSFSIFLLALLSCLTTPASA 191

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0  
Maximum DB seq length: 15

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001s:\*  
5: Geneseq2002s:\*  
6: Geneseq2003as:\*  
7: Geneseq2003bs:\*  
8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	7.9	15	2	AAR84495 Hepatitis
2	15	7.9	15	2	AAR67634 Epitope O
3	15	7.9	15	2	AAR85416 Helper T-
4	15	7.9	15	4	AAG62639 Hepatitis
5	15	7.9	15	4	AAG62640 Hepatitis
6	15	7.9	15	4	AAG62640 Hepatitis
7	15	7.9	15	4	AAG62640 Hepatitis
8	15	7.9	15	4	AAG62640 Hepatitis
9	15	7.9	15	4	AAG62640 Hepatitis
10	15	7.9	15	4	AAG62640 Hepatitis
11	15	7.9	15	4	AAG62640 Hepatitis
12	15	7.9	15	4	AAG62640 Hepatitis
13	15	7.9	15	4	AAG62640 Hepatitis
14	15	7.9	15	4	AAG62640 Hepatitis
15	15	7.9	15	4	AAG62640 Hepatitis
16	15	7.9	15	4	AAG62640 Hepatitis
17	15	7.9	15	4	AAG62640 Hepatitis
18	15	7.9	15	4	AAG62640 Hepatitis
19	15	7.9	15	4	AAG62640 Hepatitis
20	15	7.9	15	4	AAG62640 Hepatitis
21	15	7.9	15	4	AAG62640 Hepatitis
22	15	7.9	15	4	AAG62640 Hepatitis
23	15	7.9	15	4	AAG62640 Hepatitis
24	15	7.9	15	4	AAG62640 Hepatitis
25	15	7.9	15	5	AAM50818 HCV core

26	15	7.9	15	7	ABR55576	ABR55576 Amino aci
27	15	7.9	15	7	ABR55581	ABR55581 Amino aci
28	15	7.9	15	7	ABR55575	ABR55575 Amino aci
29	15	7.9	15	7	ABR55589	ABR55589 Amino aci
30	15	7.9	15	7	ABR55577	ABR55577 Amino aci
31	15	7.9	15	7	ABR55580	ABR55580 Amino aci
32	15	7.9	15	7	ADP77781	ADP77781 Hepatitis
33	15	7.9	15	7	ADH19093	ADH19093 HCV core
34	15	7.9	15	7	ADH19092	ADH19092 HCV core
35	15	7.9	15	7	ADH19096	ADH19096 HCV core
36	15	7.9	15	7	ADH19097	ADH19097 HCV core
37	15	7.9	15	7	ADH19081	ADH19081 HCV core
38	15	7.9	15	7	ADH19105	ADH19105 HCV core
39	15	7.9	15	8	AD134997	AD134997 Hepatitis
40	15	7.9	15	8	AD134996	AD134996 Hepatitis
41	15	7.9	15	8	ADL25844	ADL25844 Synthetic
42	15	7.9	15	8	ADL25845	ADL25845 Synthetic
43	15	7.9	15	8	ADL25851	ADL25851 Synthetic
44	15	7.9	15	8	ADL25860	ADL25860 Synthetic
45	15	7.9	15	8	ADL25869	ADL25869 Synthetic

ALIGNMENTS

RESULT 1  
AAR84495  
ID AAR84495 standard; peptide; 15 AA.  
XX AAR84495;  
XX  
DT 06-JAN-1997 (first entry)  
XX  
DE Hepatitis C virus peptide CORE 2 (residues 5-19).  
XX  
XX Hepatitis C virus; HCV; immunogen; core region; nucleocapsid;  
KW immunodominant; T cell epitope; vaccine.  
XX  
OS Hepatitis C virus.  
XX  
PN WO9512677-A2.  
XX  
PD 11-MAY-1995.  
XX  
PF 28-OCT-1994; 94WO-EP003555.  
XX  
PR 04-NOV-1993; 93EP-00402718.  
XX  
(INNO-) INNOGENETICS NV.  
Leroux-Roels G, Deleys R, Maertens G;  
WPI, 1995-193822/25.  
Hepatitis C Virus immunogenic polypeptide contg. a T-cell stimulating epitope - from core, E1, E2 and NS3 regions, useful in production of vaccines, therapeutic agents, etc.  
Example 4; Page 51; 105pp; English.  
A series of overlapping peptides (including the present sequence) was synthesised based on sequences in the core, E1 and E2/NS1 regions of hepatitis C virus. The peptides were used as antigens in lympho-proliferative assays to identify the main T-cell epitopes

Query Match

Best Local Similarity 7.9%; Score 15; DB 2; Length 15;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PKPQKTKRNTNRRP 19

|||||

Dd 1 PKPQRTKXNTNRRP 15

RESULT 2  
AAR67634  
ID AAR67634 standard; protein; 15 AA.  
XX AC  
XX AAW85416;  
DT 16-FEB-1999 (first entry)  
XX DE  
XX Helper T-cell class II peptide derived from core protein.  
XX KW  
XX Helper T-cell peptide; human leucocyte antigen; HLA; DR4w4; DRI; DR7;  
KW cytotoxic T lymphocyte; CTL; hepatitis; autoimmune disease;  
KW acquired immune deficiency syndrome; malaria; cancer;  
KW allograft rejection; allergy; Lyme disease; hepatitis;  
KW post-streptococcal endocarditis; glomerulonephritis;  
KW food hypersensitivity.  
XX OS  
XX Synthetic.  
OS Hepatitis C virus.  
XX PN  
XX WO9832456-A1.  
XX PD  
XX 30-JUL-1998.  
XX PF  
XX 23-JAN-1998; 98WO-US001373.  
PR XX  
XX 23-JAN-1997; 97US-0036713P.  
PR 07-FEB-1997; 97US-0037432P.  
XX PA  
XX (EPIM-) EPIMUNE INC.  
XX PI  
XX Sette A, Sidney J, Southwood S;  
DR WPI; 1998-427679/36.  
XX PT  
XX Composition containing peptide that induces cytotoxic T lymphocyte response, and helper peptide - can bind to human leucocyte antigen alleles, used to treat or prevent cancers, parasitic infections and autoimmune disease.  
XX PS  
XX Disclosure; Page 42; 51pp; English.  
XX CC  
XX AAW85284-451 represent helper T-cell class II peptides, which can bind to the human leucocyte antigens (HLA) DR4w4, DRI and DR7. The peptides are used in the course of the invention. The specification describes peptides that induce a cytotoxic T lymphocyte (CTL) response, and T-helper peptides, that are used together to generate a CTL response for the treatment or prevention of viral, fungal, bacterial or parasitic infections (e.g. hepatitis, acquired immune deficiency syndrome or malaria) or cancer (e.g. renal or cervical carcinoma, lymphoma, prostate cancer or condyloma acuminatum). Helper T-cell peptides may be used alone to induce a helper T cell response, e.g. in cases of autoimmune disease, allograft rejection, allergy, Lyme disease, hepatitis, post-streptococcal endocarditis, glomerulonephritis and food hypersensitivity  
XX SQ  
XX Sequence 15 AA;  
  
Query Match 7.9%; Score 15; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 34 VYLLPRRGRLGVRA 48  
|||  
Db 1 VYLLPRRGRLGVRA 15  
  
RESULT 4  
AAG62639  
ID AAG62639 standard; peptide; 15 AA.  
XX AC  
XX AAG62639;  
XX DT  
XX 11-SEP-2001 (first entry)  
XX DE  
XX Hepatitis C virus core protein epitope #1.

Dd 1 PKPQRTKXNTNRRP 15

RESULT 2  
AAR67634  
ID AAR67634 standard; protein; 15 AA.  
XX AC  
XX AAR67634;  
XX DT  
XX 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 18-AUG-1995 (first entry)  
XX DE  
XX Epitope of Non-A Non-B hepatitis virus structural protein.  
DE DE  
XX Non-A Non-B hepatitis virus; structural region; cDNA to genomic RNA;  
KW detection; reagent; anti-Non-A Non-B hepatitis virus antibody; vaccine;  
KW antigen; epitope; diagnosis.  
XX OS  
XX Non-A.  
OS non-B hepatitis virus.  
OS OS  
PN EF628572-A2.  
XX PD  
XX 14-DEC-1994.  
XX PF  
XX 27-MAY-1994; 94BP-00108256.  
XX PR  
XX 28-MAY-1993; 93JP-00126709.  
PR 02-MAR-1994; 94JP-00032201.  
XX PA  
XX (EISA ) EISAI CO LTD.  
XX PI  
XX Aoyama M, Obara T, Tohmatsu J, Sawada T, Hosoda T, Iwasaki Y;  
PI Arima T;  
XX DR  
XX WPI; 1995-015655/03.  
XX PT  
XX New non-A non-B hepatitis virus sub-type - used to develop prods. for detection, diagnosis, prevention and treatment of non-A non-B hepatitis.  
XX PS  
XX Example 3; Page 39; 59pp; English.  
XX CC  
XX AAR67634 is designated M2ae and is located at positions 39-53 of the structural protein (AAR67617) of the Non-A Non-B (NANB) hepatitis virus encoded by a partial cDNA to genomic RNA sequence. Regions of the polypeptide were studied for suitability as an epitope. The positive ratios of the peptide epitope to the plasma of donors, who had been judged as positive with a reagent of the second generation, i.e., a second generational anti hepatitis virus antibody detection reagent, were determined. This peptide had a positive ratio (%) of 62.3. The novel epitopes are effective in the diagnosis of NANB hepatitis. The nucleotide sequences (see AQO5817-19) were isolated from the plasma of donors in Japan with high s-GTP levels, and were found to be different from previously reported NANB hepatitis viruses. The DNA can be used as a reagent for detecting the NANB hepatitis viral gene. The polypeptides can be used as reagents for detecting anti-NANB hepatitis antibodies or as a NANB hepatitis viral vaccine. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS field.)  
XX SQ  
XX Sequence 15 AA;  
  
Query Match 7.9%; Score 15; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 39 RRGFRGLGVRATRKTS 53  
|||||  
Db 1 RRGFRGLGVRATRKTS 15  
  
RESULT 3  
AAW85416

XX Hepatitis C virus; HCV; vaccine; charged antigen; charged carrier;  
KW electrostatic association; core protein; epitope.  
XX Hepatitis C virus.  
XX WO200137869-A1.  
XX 31-MAY-2001.  
XX 17-NOV-2000; 2000WO-AU001410.  
XX 19-NOV-1999; 99US-0166652P.  
XX 11-AUG-2000; 2000US-0224362P.  
XX (CSLC-) CSL LTD.  
XX (CHIR ) CHIRON CORP.  
XX Drane D, Cox J, Houghton M, Palliard X;  
XX WPI; 2001-367618/38.  
XX Novel immunogenic complex, useful in vaccine compositions for therapeutic  
PT and prophylactic treatment of disease conditions resulting from hepatitis  
PT C virus infection, has charged organic carrier and charged antigen.  
XX Example 4; Page 46; 67pp; English.  
XX The present invention relates to an immunogenic complex, comprising an  
CC electrostatically associated charged organic carrier and a charged  
CC antigen, where the antigen is a protein from the hepatitis C virus (HCV).  
CC This can be used in a vaccine to elicit an immune response against HCV.  
CC and therefore prevent infection by the virus. The present sequence is a  
CC HCV core protein derived epitope described in the exemplification of the  
CC invention  
XX Sequence 15 AA;  
XX Query Match 7.9%; Score 15; DB 4; Length 15;  
XX Best Local Similarity 100.0%; Pred. No. 2.6e-06; Mismatches 0; Indels 0; Gaps 0;  
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 121 KVIDTLTCGFADLMG 135  
DB 1 KVIDTLTCGFADLMG 15  
RESULT 5  
AAG62640  
ID AAG62640 standard; peptide; 15 AA.  
XX AAG62640;  
XX 11-SEP-2001 (first entry)  
XX Hepatitis C virus core protein epitope #2.  
XX Hepatitis C virus; HCV; vaccine; charged antigen; charged carrier;  
KW electrostatic association; core protein; epitope.  
XX Hepatitis C virus.  
XX WO200137869-A1.  
XX 31-MAY-2001.  
XX 17-NOV-2000; 2000WO-AU001410.  
XX 19-NOV-1999; 99US-0166652P.  
XX 11-AUG-2000; 2000US-0224362P.  
XX (CSLC-) CSL LTD.  
XX (CHIR ) CHIRON CORP.  
XX Drane D, Cox J, Houghton M, Palliard X;  
XX WPI; 2001-367618/38.  
XX Novel immunogenic complex, useful in vaccine compositions for therapeutic  
PT and prophylactic treatment of disease conditions resulting from hepatitis  
PT C virus infection, has charged organic carrier and charged antigen.  
XX Example 4; Page 46; 67pp; English.  
XX The present invention relates to an immunogenic complex, comprising an  
CC electrostatically associated charged organic carrier and a charged  
CC antigen, where the antigen is a protein from the hepatitis C virus (HCV).  
CC This can be used in a vaccine to elicit an immune response against HCV.  
CC and therefore prevent infection by the virus. The present sequence is a  
CC HCV core protein derived epitope described in the exemplification of the  
CC invention  
XX Sequence 15 AA;  
XX Query Match 7.9%; Score 15; DB 4; Length 15;  
XX Best Local Similarity 100.0%; Pred. No. 2.6e-06; Mismatches 0; Indels 0; Gaps 0;  
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 121 KVIDTLTCGFADLMG 135  
DB 1 KVIDTLTCGFADLMG 15  
RESULT 5  
AAG62640  
ID AAG62640 standard; peptide; 15 AA.  
XX AAG62640;  
XX 11-SEP-2001 (first entry)  
XX Hepatitis C virus core protein epitope #2.  
XX Hepatitis C virus; HCV; vaccine; charged antigen; charged carrier;  
KW electrostatic association; core protein; epitope.  
XX Hepatitis C virus.  
XX WO200137869-A1.  
XX 31-MAY-2001.  
XX 17-NOV-2000; 2000WO-AU001410.  
XX 19-NOV-1999; 99US-0166652P.  
XX 11-AUG-2000; 2000US-0224362P.  
XX (CSLC-) CSL LTD.  
XX (CHIR ) CHIRON CORP.

XX Drane D, Cox J, Houghton M, Palliard X;  
XX WPI; 2001-367618/38.  
XX Novel immunogenic complex, useful in vaccine compositions for therapeutic  
PT and prophylactic treatment of disease conditions resulting from hepatitis  
PT C virus infection, has charged organic carrier and charged antigen.  
XX Example 4; Page 46; 67pp; English.  
XX The present invention relates to an immunogenic complex, comprising an  
CC electrostatically associated charged organic carrier and a charged  
CC antigen, where the antigen is a protein from the hepatitis C virus (HCV).  
CC This can be used in a vaccine to elicit an immune response against HCV.  
CC and therefore prevent infection by the virus. The present sequence is a  
CC HCV core protein derived epitope described in the exemplification of the  
CC invention  
XX Sequence 15 AA;  
XX Query Match 7.9%; Score 15; DB 4; Length 15;  
XX Best Local Similarity 100.0%; Pred. No. 2.6e-06; Mismatches 0; Indels 0; Gaps 0;  
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 86 YGNEGCGWAGWLLSP 100  
DB 1 YGNEGCGWAGWLLSP 15  
RESULT 6  
AAJ03729  
ID AAJ03729 standard; peptide; 15 AA.  
XX AAJ03729;  
XX 02-JUL-2001 (first entry)  
XX Hepatitis C virus epitope #3720.  
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;  
KW antiviral.  
XX Hepatitis C virus.  
XX WO200121189-A1.  
XX 29-MAR-2001.  
XX 19-JUL-2000; 2000WO-US019774.  
XX 19-JUL-1999; 99US-00357737.  
XX (EPIM-) EPIMMUNE INC.  
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesenut R;  
XX Baker DM, Cellis E, Kubo RT, Grey HM;  
XX WPI; 2001-308046/32.  
XX A new composition useful as a vaccine against hepatitis C virus.  
XX Disclosure; Page 182; 214pp; English.  
XX The present invention describes a composition comprising a prepared  
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.  
CC These are derived from HCV HLA-binding motifs. They are useful in  
CC vaccines for the prevention and treatment of HCV infection in humans. The  
CC present sequence is an epitope used in the disclosure of the invention  
XX Sequence 15 AA;  
XX Query Match 7.9%; Score 15; DB 4; Length 15;

```

Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 RLGVTRATKTSRSQ 57
DB 1 RLGVTRATKTSRSQ 15

RESULT 7
AAJ03568
ID AAJ03568 standard; peptide; 15 AA.
XX
AC AAJ03568;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #3559.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
SQ Sequence 15 AA;

Query Match 7.9%; Score 15; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 LGKVIDTLTCGFADL 133
DB 1 LGKVIDTLTCGFADL 15

RESULT 9
AAJ03352
ID AAJ03352 standard; peptide; 15 AA.
XX
AC AAJ03352;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #3343.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
SQ Sequence 15 AA;

Query Match 7.9%; Score 15; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 GCGWAGWLLSPRGR 104
DB 1 GCGWAGWLLSPRGR 15

RESULT 8
AAJ03200
ID AAJ03200 standard; peptide; 15 AA.
XX
AC AAJ03200;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #3191.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
antiviral.
XX
OS Hepatitis C virus.

```

The present invention describes a composition comprising a prepared hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.

CC These are derived from HCV HLA-binding motifs. They are useful in  
CC vaccines for the prevention and treatment of HCV infection in humans. The  
CC present sequence is an epitope used in the disclosure of the invention  
XX  
XX Sequence 15 AA;

Query Match 7.9%; Score 15; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 GCSFSIFLLALLSCL 185  
|||||  
Db 1 GCSFSIFLLALLSCL 15

RESULT 10  
AAJ03468  
ID AAJ03468 standard; peptide; 15 AA.

XX AC AAJ03468;

DT 02-JUL-2001 (first entry)

XX Hepatitis C virus epitope #3459.

XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;  
KW antiviral.

XX OS Hepatitis C virus.

XX PN WO200121189-A1.

XX PD 29-MAR-2001.

XX PF 19-JUL-2000; 2000WO-US019774.

XX PR 19-JUL-1999; 99US-00357737.

XX PA (EPIM-) EPIMMUNE INC.

PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
PI Baker DM, Celis E, Kubo RT, Grey HM;

XX WPI; 2001-308046/32.

XX A new composition useful as a vaccine against hepatitis C virus.

XX PS Disclosure; Page 178; 214pp; English.

XX The present invention describes a composition comprising a prepared  
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.  
CC These are derived from HCV HLA-binding motifs. They are useful in  
CC vaccines for the prevention and treatment of HCV infection in humans. The  
CC present sequence is an epitope used in the disclosure of the invention

XX Sequence 15 AA;

Query Match 7.9%; Score 15; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 IDTLTGFADLMGYI 137  
|||||  
Db 1 IDTLTGFADLMGYI 15

RESULT 11  
AAJ03046  
ID AAJ03046 standard; peptide; 15 AA.

XX AC AAJ03046;

DT 02-JUL-2001 (first entry)

XX Hepatitis C virus epitope #3037.  
DE  
XX  
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;  
KW antiviral.  
XX  
XX Hepatitis C virus.  
OS  
XX WO200121189-A1.  
PN  
XX 29-MAR-2001.  
PD  
XX 19-JUL-2000; 2000WO-US019774.  
PF  
XX 19-JUL-1999; 99US-00357737.  
PR  
XX (EPIM-) EPIMMUNE INC.  
PA  
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
PI Baker DM, Celis E, Kubo RT, Grey HM;  
XX WPI; 2001-308046/32.  
DR  
XX A new composition useful as a vaccine against hepatitis C virus.  
PT  
XX Disclosure; Page 174; 214pp; English.  
PS  
XX The present invention describes a composition comprising a prepared  
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.  
CC These are derived from HCV HLA-binding motifs. They are useful in  
CC vaccines for the prevention and treatment of HCV infection in humans. The  
CC present sequence is an epitope used in the disclosure of the invention

XX Sequence 15 AA;  
Query Match 7.9%; Score 15; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 GCSFSIFLLALLSCL 185  
|||||  
Db 1 GCSFSIFLLALLSCL 15

RESULT 12  
AAJ03983  
ID AAJ03983 standard; peptide; 15 AA.

XX AC AAJ03983;

XX DT 02-JUL-2001 (first entry)

XX Hepatitis C virus epitope #3974.

XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;  
KW antiviral.

XX OS Hepatitis C virus.

XX PN WO200121189-A1.

XX PD 29-MAR-2001.

XX PF 19-JUL-2000; 2000WO-US019774.

XX PR 19-JUL-1999; 99US-00357737.

XX PA (EPIM-) EPIMMUNE INC.

XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
PI Baker DM, Celis E, Kubo RT, Grey HM;

XX WPI; 2001-308046/32.

```
XX A new composition useful as a vaccines against hepatitis C virus.
PT Example 5; Page 198; 214pp; English.
XX
XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX Sequence 15 AA;
SQ
Query Match 7.9%; Score 15; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 119 LGKVIDTLTCGFADL 133
Db 1 LGKVIDTLTCGFADL 15
RESULT 13
AAJ03360
ID AAJ03360 standard; peptide; 15 AA.
XX
AC AAJ03360;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #3351.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
PT A new composition useful as a vaccines against hepatitis C virus.
XX
PS Disclosure; Page 177; 214pp; English.
XX
CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX Sequence 15 AA;
SQ
Query Match 7.9%; Score 15; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 120 GKVIDTLTCGFADLM 134
Db 1 GKVIDTLTCGFADLM 15
```

```
RESULT 14
AAJ03151
ID AAJ03151 standard; peptide; 15 AA.
XX
AC AAJ03151;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #3142.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
PT A new composition useful as a vaccines against hepatitis C virus.
XX
PS Disclosure; Page 175; 214pp; English.
XX
CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX Sequence 15 AA;
SQ
Query Match 7.9%; Score 15; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 34 VYLLPRRGPRGLGVRA 48
Db 1 VYLLPRRGPRGLGVRA 15
RESULT 15
AAJ03507
ID AAJ03507 standard; peptide; 15 AA.
XX
AC AAJ03507;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #3498.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
```

XX (EPIM-) EPIMUNE INC.  
 XX PA Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
 XX PI Baker DM, Celis E, Kubo RT, Grey HM;  
 XX PI WPI: 2001-308046/32.  
 XX DR  
 XX PT A new composition useful as a vaccines against hepatitis C virus.  
 XX PS Disclosure; Page 178; 214pp; English.  
 XX CC The present invention describes a composition comprising a prepared  
 CC hepatitis C virus (HCV) epitope such as those given in AAJ0010-AAJ04121.  
 CC These are derived from HCV HLA-binding motifs. They are useful in  
 CC vaccines for the prevention and treatment of HCV infection in humans. The  
 CC present sequence is an epitope used in the disclosure of the invention  
 XX SQ Sequence 15 AA;  
 SQ  
 Query Match 7.9%; Score 15; DB 4; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 119 LGKVIDTLTCGFADL 133  
 DB 1 LGKVIDTLTCGFADL 15

Search completed: July 15, 2005, 10:23:34  
 Job time : 161 secs

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2005, 10:18:06 ; Search time 42 Seconds  
(without alignments)  
339.476 Million cell updates/sec

Title: US-09-084-691B-206  
Perfect score: 191  
Sequence: 1 MSTLPKPKYKRNTRPT.....CSFSIFLLALLSCLTTPASA 191

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 164526

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	7.9	15	4	US-09-009-953-243
2	15	7.9	15	4	US-08-635-886C-1
3	15	7.9	15	4	US-08-974-690C-1
4	15	7.9	15	4	US-09-790-497A-134
5	15	7.9	15	4	US-09-576-824A-134
6	15	7.9	15	4	US-08-974-685-1
7	14	7.3	15	3	US-08-604-365-15
8	14	7.3	15	4	US-09-689-678-15
9	13	6.8	15	4	US-09-009-953-180
10	12	6.3	12	1	US-08-102-738-22
11	12	6.3	12	1	US-08-102-738-23
12	12	6.3	12	1	US-08-102-738-24
13	11	5.8	11	2	US-08-737-085A-12
14	11	5.8	11	2	US-08-501-195-3
15	11	5.8	11	2	US-08-466-975A-3
16	11	5.8	11	2	US-08-391-671A-3
17	11	5.8	11	2	US-08-467-902A-3
18	11	5.8	11	3	US-08-159-339A-1139
19	11	5.8	11	3	US-09-246-258-12
20	11	5.8	11	3	US-09-532-106-12
21	11	5.8	11	3	US-09-275-265-3
22	11	5.8	11	4	US-09-839-666-12
23	11	5.8	11	4	US-09-311-784A-284
24	11	5.8	11	4	US-09-941-611-3
25	11	5.8	11	4	US-09-790-497A-39
26	11	5.8	12	1	US-08-102-738-17
27	11	5.8	12	3	US-08-604-365-19

28 11 5.8 12 4 US-09-689-678-19 Sequence 19, Appl  
29 11 5.8 13 1 US-07-681-701-5 Sequence 5, Appl  
30 11 5.8 13 2 US-08-146-028-39 Sequence 39, Appl  
31 11 5.8 13 3 US-08-723-425A-39 Sequence 39, Appl  
32 11 5.8 13 3 US-09-112-206-39 Sequence 39, Appl  
33 11 5.8 13 4 US-09-576-824A-39 Sequence 39, Appl  
34 11 5.8 13 4 US-09-680-497-39 Sequence 52, Appl  
35 10 5.2 10 4 US-08-197-484-52 Sequence 131, Appl  
36 10 5.2 10 4 US-08-197-484-131 Sequence 131, Appl  
37 10 5.2 10 4 US-09-311-784A-262 Sequence 262, Appl  
38 10 5.2 10 4 US-10-133-007-4 Sequence 4, Appl  
39 10 5.2 10 5 PCT-US95-02121-52 Sequence 52, Appl  
40 10 5.2 10 5 PCT-US95-02121-131 Sequence 131, Appl  
41 10 5.2 11 1 US-08-102-738-21 Sequence 21, Appl  
42 10 5.2 11 4 US-09-576-824A-514 Sequence 514, Appl  
43 10 5.2 12 3 US-08-604-365-17 Sequence 17, Appl  
44 10 5.2 12 4 US-09-689-678-17 Sequence 17, Appl  
45 9 4.7 9 2 US-08-146-028-182 Sequence 182, Appl

#### ALIGNMENTS

RESULT 1  
US-09-009-953-243  
; Sequence 243, Application US/09009953  
; Patent No. 6413517  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; TITLE OF INVENTION: Identification of Broadly  
; REACTIVE DR RESTRICTED EPITOPES  
; NUMBER OF SEQUENCES: 274  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/009,953  
; FILING DATE: 21-Jan-1998  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/036,713  
; FILING DATE: 23-JAN-1997  
; APPLICATION NUMBER: US 60/037,432  
; FILING DATE: 07-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen Lauver  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 018623-01152005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-576-0200  
; TELEFAX: 415-576-0300  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 243:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 243:  
US-09-009-953-243

Query Match 7.9%; Score 15; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 VYLLPRGPRLGURA 48  
|||||  
Db 1 VYLLPRGPRLGURA 15

## RESULT 2

US-08-635-886C-1  
; Sequence 1, Application US/08635886C  
; Patent No. 6355114  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, Geert  
; APPLICANT: DELEYS, Robert  
; APPLICANT: MAERTENS, Geert  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
; FILE REFERENCE: 2752-18  
; CURRENT APPLICATION NUMBER: US/08/635,886C  
; PRIOR FILING DATE: 1996-04-25  
; PRIOR FILING DATE: 1994-10-28  
; PRIOR APPLICATION NUMBER: PCT/EP94/03555  
; PRIOR FILING DATE: 1993-11-04  
; PRIOR FILING DATE: 1993-11-04  
; NUMBER OF SEQ ID NOS: 286  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
US-08-635-886C-1

Query Match 7.9%; Score 15; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PKPQKTKRNTNRRP 19  
|||||  
Db 1 PKPQKTKRNTNRRP 15

## RESULT 3

US-08-974-690C-1  
; Sequence 1, Application US/08974690C  
; Patent No. 661333  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, Geert  
; APPLICANT: DELEYS, Robert  
; APPLICANT: MAERTENS, Geert  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
; FILE REFERENCE: 2551-94  
; CURRENT APPLICATION NUMBER: US/08/974,690C  
; CURRENT FILING DATE: 1997-11-19  
; PRIOR APPLICATION NUMBER: PCT/EP94/03555  
; PRIOR FILING DATE: 1994-10-28  
; PRIOR APPLICATION NUMBER: EP 93402718.6  
; PRIOR FILING DATE: 1993-11-04  
; NUMBER OF SEQ ID NOS: 286  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
US-08-974-690C-1

Query Match 7.9%; Score 15; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PKPQKTKRNTNRRP 19  
|||||  
Db 1 PKPQKTKRNTNRRP 15

## RESULT 4

US-09-790-497A-134  
; Sequence 134, Application US/09790497A  
; Patent No. 6649735  
; GENERAL INFORMATION:  
; APPLICANT: De Leys, Robert  
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING  
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN  
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF  
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT  
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS  
; TITLE OF INVENTION: CONTAINING THEM

FILE REFERENCE: 2752-16  
CURRENT APPLICATION NUMBER: US/09/790,497A  
CURRENT FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: 09/576,824  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 08/723,425  
PRIOR FILING DATE: 1996-09-30  
PRIOR APPLICATION NUMBER: 09/146,028  
PRIOR FILING DATE: 1993-11-22  
PRIOR APPLICATION NUMBER: PCT/EP93/00517  
PRIOR FILING DATE: 1993-03-08  
PRIOR APPLICATION NUMBER: EP 92400598.6  
PRIOR FILING DATE: 1992-03-06  
NUMBER OF SEQ ID NOS: 600  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 134  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Hepatitis C virus  
US-09-790-497A-134

Query Match 7.9%; Score 15; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PKPQKTKRNTNRRP 19  
|||||  
Db 1 PKPQKTKRNTNRRP 15

## RESULT 5

US-09-576-824A-134  
; Sequence 134, Application US/09576824A  
; Patent No. 6667387  
; GENERAL INFORMATION:  
; APPLICANT: De Leys, Robert  
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING  
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN  
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF  
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT  
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS  
; TITLE OF INVENTION: CONTAINING THEM

FILE REFERENCE: 2752-11  
CURRENT APPLICATION NUMBER: US/09/576,824A  
CURRENT FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 08/723,425  
PRIOR FILING DATE: 1996-09-30  
PRIOR APPLICATION NUMBER: 09/146,028  
PRIOR FILING DATE: 1993-11-22  
PRIOR APPLICATION NUMBER: PCT/EP93/00517  
PRIOR FILING DATE: 1993-03-08  
PRIOR APPLICATION NUMBER: EP 92400598.6  
PRIOR FILING DATE: 1992-03-06  
NUMBER OF SEQ ID NOS: 600  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 134  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Hepatitis C virus

QY 5 PKPQKTKRNTNRRP 19  
|||||  
Db 1 PKPQKTKRNTNRRP 15



```

;
; COUNTRY: USA
; ZIP: 10022
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/689,678
; FILING DATE: 13-Oct-2000
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/604,365
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 07/977,398
; FILING DATE: 11-MARCH-1993
; APPLICATION NUMBER: PCT/EP92/01468
; FILING DATE: 30-JUNE-1992
; APPLICATION NUMBER: DE 41 22 160.5
; FILING DATE: 04-JULY-1991
; APPLICATION NUMBER: DE 41 41 304.0
; FILING DATE: 14-DEC-1991
; APPLICATION NUMBER: DE 42 09 216.9
; FILING DATE: 21-MARCH-1992
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: BOER 1010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
;
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-689-678-15
;
; Query Match 7.3%; Score 14; DB 4; Length 15;
; Best Local Similarity 100.0%; Pred. No. 2e-06;
; Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 5 PKPQKTKRNTNR 18
; DB 2 PKPQKTKRNTNR 15
;
; RESULT 9
; US-09-009-953-180
; Sequence 180, Application US/09009953
; Patent No. 6413517
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; TITLE OF INVENTION: Identification of Broadly
; REACTIVE DR RESTRICTED EPITOPES
;
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,953
; FILING DATE: 21-Jan-1998

```

```

;
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,713
; FILING DATE: 23-JAN-1997
; APPLICATION NUMBER: US 60/037,432
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 018623-011520US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; TELEX: <Unknown>
;
; INFORMATION FOR SEQ ID NO: 180:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 180:
US-09-009-953-180
;
; Query Match 6.8%; Score 13; DB 4; Length 15;
; Best Local Similarity 100.0%; Pred. No. 1.7e-05;
; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 163 NYATGNLPGCSFS 175
; DB 3 NYATGNLPGCSFS 15
;
; RESULT 10
; US-08-102-738-22
; Sequence 22, Application US/08102738
; Patent No. 5674676
; GENERAL INFORMATION:
; APPLICANT: Seidel, Christoph; Wienhues, Ursula-Henrike;
; APPLICANT: Bayer, Hubert; Jung, Guenther-Gerhard; Ihlenfeldt, Hans-
; APPLICANT: Georg
; TITLE OF INVENTION: HCV Peptide Antigens And Method Of
; DETERMINING HCV
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/102,738
; FILING DATE: 6-AUG-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 42 26 093.0
; FILING DATE: 7-AUG-1992
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 42 40 980.2
; FILING DATE: 5-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: BOER 1019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884

```

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-102-738-22

Query Match 6.3%; Score 12; DB 1; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.00013;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 GOVGGVLLPR 39

|||||

Db 1 GOVGGVLLPR 12

RESULT 11

US-08-102-738-23

Sequence 23, Application US/08102738

Patent No. 5674676

GENERAL INFORMATION:

APPLICANT: Seidel, Christoph; Wienhues, Ursula-Henrike;

APPLICANT: Bayer, Hubert; Jung, Guenther-Gerhard; Ihlenfeldt, Hans-

APPLICANT: Georg

TITLE OF INVENTION: HCV Peptide Antigens And Method Of

TITLE OF INVENTION: Determining HCV

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/102,738

FILING DATE: 6-AUG-1993

CLASSIFICATION: 424

PRIOR APPLICATION NUMBER:

APPLICATION NUMBER: DE 42 26 093.0

FILING DATE: 7-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 42 40 980.2

FILING DATE: 5-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Tsai, Christine H.

REGISTRATION NUMBER: 34,266

REFERENCE/DOCKET NUMBER: BOER 1019

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-102-738-23

Query Match

Best Local Similarity 6.3%; Score 12; DB 1; Length 12;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 VCGVLLPRGP 42

|||||

Db 1 VCGVLLPRGP 12

RESULT 12

US-08-102-738-24

Sequence 24, Application US/08102738

Patent No. 5674676

GENERAL INFORMATION:

APPLICANT: Seidel, Christoph; Wienhues, Ursula-Henrike;

APPLICANT: Bayer, Hubert; Jung, Guenther-Gerhard; Ihlenfeldt, Hans-

APPLICANT: Georg

TITLE OF INVENTION: HCV Peptide Antigens And Method Of

TITLE OF INVENTION: Determining HCV

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/102,738

FILING DATE: 6-AUG-1993

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 42 26 093.0

FILING DATE: 7-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 42 40 980.2

FILING DATE: 5-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Tsai, Christine H.

REGISTRATION NUMBER: 34,266

REFERENCE/DOCKET NUMBER: BOER 1019

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-102-738-24

Query Match

Best Local Similarity 6.3%; Score 12; DB 1; Length 12;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 VYLLPRGRPLG 45

|||||

Db 1 VYLLPRGRPLG 12

RESULT 13

US-08-737-085A-12

Sequence 12, Application US/08737085A

Patent No. 5869232

GENERAL INFORMATION:

APPLICANT: SALLBERG, MATTI

TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY

TITLE OF INVENTION: EXCHANGER

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: DARBY & DARBY PC

STREET: 805 Third Avenue

;  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/737,085A  
; FILING DATE: 27-DEC-1996  
; CLASSIFICATION: 426  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Green, Reza  
; REGISTRATION NUMBER: 38,475  
; REFERENCE/DOCKET NUMBER: 3846/0C569  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-527-7659  
; TELEFAX: 212-753-6237  
; TELEX: 236687  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-737-085A-12

Query Match 5.8%; Score 11; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ORKTKRNTNR 18  
Db 1 ORKTKRNTNR 11

RESULT 14  
US-08-501-195-3  
; Sequence 3, Application US/085011195  
; Patent No. 5871904  
; GENERAL INFORMATION:  
; APPLICANT: KASHIWAKUMA, Tomiko  
; APPLICANT: YAGI, Shintaro  
; APPLICANT: HASEGAWA, Akira  
; APPLICANT: KAJITA, Tadashi  
; APPLICANT: OTA, Yosuke  
; APPLICANT: MORI, Hiroyuki  
; TITLE OF INVENTION: IMMUNOASSAY OF NON-A, NON-B  
; TITLE OF INVENTION: HEPATITIS VIRUS-RELATED ANTIGENS, MONOCLONAL  
; TITLE OF INVENTION: ANTIBODIES FOR USE THEREIN, AND HYBRIDOMAS  
; TITLE OF INVENTION: PRODUCING THE ANTIBODIES  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: KAWAGUTI & PARTNERS  
; STREET: Yamada Bldg., 1-14  
; CITY: Shinjuku 1-chome, Shinjuku-ku  
; STATE: Tokyo  
; COUNTRY: Japan  
; ZIP: 160  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh SE, Text file  
; SOFTWARE: Microsoft Word Version #4.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/501,195  
; FILING DATE:

;  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-183904  
; FILING DATE: July 12, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME:  
; REGISTRATION NUMBER:  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE:  
; TELEFAX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: NON-A, NON-B hepatitis virus  
; US-08-501-195-3

Query Match 5.8%; Score 11; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 PGGGQIVGGVY 35  
Db 1 PGGGQIVGGVY 11

RESULT 15  
US-08-466-975A-3  
; Sequence 3, Application US/08466975A  
; Patent No. 5910404  
; GENERAL INFORMATION:  
; APPLICANT: DELEYS, ROBERT J  
; APPLICANT: POLLET, DIRK  
; APPLICANT: MAERTENS, GEERT  
; APPLICANT: VAN HEUVERSWUN, HUGO  
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF  
; TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHUYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22201  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/466,975A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/391,671  
; FILING DATE:  
; APPLICATION NUMBER: US 07/920,286  
; FILING DATE: 14-OCT-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/EP91/02409  
; FILING DATE: 13-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 90124241.2  
; FILING DATE: 14-DEC-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SADOFF, B.J.  
; REGISTRATION NUMBER: 36,663

```

; REFERENCE/DOCKET NUMBER: 1487-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-466-975A-3

```

```

Query Match      5.8%; Score 11; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      8 ORKTKNTNRR 18
Db      1 ORKTKNTNRR 11

```

Search completed: July 15, 2005, 10:28:09  
Job time : 43 secs

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2005, 10:26:41 ; Search time 161 Seconds  
(without alignments)  
458.986 Million cell updates/sec

Title: US-09-084-691B-206  
Perfect score: 191  
Sequence: 1 MSTLPKPKQKTKRNTNRPT.....CSFSIFLLALLSCLTTPASA 191

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1729182 seqs, 386893608 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 287330

Minimum DB seq length: 0  
Maximum DB seq length: 15

Post-processing: Listing first 45 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
  - 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
  - 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*
  - 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep:\*
  - 18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
  - 19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep:\*
  - 20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*
  - 21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
  - 22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	7.9	15	13	US-10-103-395-243
2	15	7.9	15	14	US-10-268-561-4
3	15	7.9	15	14	US-10-268-561-5
4	15	7.9	15	14	US-10-268-561-6
5	15	7.9	15	14	US-10-268-561-9
6	15	7.9	15	14	US-10-268-561-10
7	15	7.9	15	14	US-10-268-561-18
8	15	7.9	15	14	US-10-268-569-4
9	15	7.9	15	14	US-10-268-569-5
10	15	7.9	15	14	US-10-268-569-6
11	15	7.9	15	14	US-10-268-569-9

12	15	7.9	15	14	US-10-268-569-10	Sequence 10, Appl
13	15	7.9	15	14	US-10-268-569-18	Sequence 18, Appl
14	15	7.9	15	15	US-10-651-165-1	Sequence 1, Appl
15	15	7.9	15	17	US-10-621-675-134	Sequence 134, App
16	14	7.3	14	14	US-10-211-088-243	Sequence 243, App
17	14	7.3	14	17	US-10-482-029-372	Sequence 372, App
18	14	7.3	15	14	US-10-288-561-3	Sequence 3, Appl
19	14	7.3	15	14	US-10-268-569-3	Sequence 3, Appl
20	14	7.3	15	14	US-10-371-540-15	Sequence 15, Appl
21	13	6.8	15	13	US-10-103-395-180	Sequence 180, App
22	13	6.8	15	14	US-10-268-561-14	Sequence 14, Appl
23	13	6.8	15	14	US-10-268-569-14	Sequence 14, Appl
24	12	6.3	15	14	US-10-288-561-7	Sequence 7, Appl
25	12	6.3	15	14	US-10-268-569-7	Sequence 7, Appl
26	11	5.8	11	9	US-09-839-666-12	Sequence 12, Appl
27	11	5.8	11	9	US-09-941-611-3	Sequence 3, Appl
28	11	5.8	11	9	US-09-894-018-232	Sequence 232, App
29	11	5.8	11	14	US-10-044-395-3	Sequence 3, Appl
30	11	5.8	11	14	US-10-234-579-12	Sequence 12, Appl
31	11	5.8	11	15	US-10-371-525-284	Sequence 284, App
32	11	5.8	11	15	US-10-371-069-284	Sequence 284, App
33	11	5.8	11	15	US-10-371-645-284	Sequence 284, App
34	11	5.8	11	15	US-10-371-360-284	Sequence 284, App
35	11	5.8	11	15	US-10-372-735-71	Sequence 71, Appl
36	11	5.8	11	16	US-10-474-960A-232	Sequence 232, App
37	11	5.8	11	17	US-10-822-871-3	Sequence 3, Appl
38	11	5.8	12	14	US-10-371-540-19	Sequence 19, Appl
39	11	5.8	13	17	US-10-621-675-39	Sequence 39, Appl
40	11	5.8	15	14	US-10-268-561-13	Sequence 13, Appl
41	11	5.8	15	14	US-10-268-561-19	Sequence 19, Appl
42	11	5.8	15	14	US-10-268-569-13	Sequence 13, Appl
43	10	5.2	10	8	US-08-344-824-178	Sequence 178, App
44	10	5.2	10	8	US-08-344-824-180	Sequence 180, App
45	10	5.2	10	9	US-09-894-018-217	Sequence 217, App

ALIGNMENTS

RESULT 1  
US-10-103-395-243  
; Sequence 243, Application US/10103395  
; Publication No. US20020160019A1  
; GENERAL INFORMATION:  
; APPLICANT: EPIMUNE, Inc.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; TITLE OF INVENTION: IDENTIFICATION OF BROADLY REACTIVE DR  
; FILE REFERENCE: 39963-20016.01  
; CURRENT APPLICATION NUMBER: US/10/103,395  
; CURRENT FILING DATE: 2003-01-03  
; PRIOR APPLICATION NUMBER: US 09/009,953  
; PRIOR FILING DATE: 1998-01-21  
; PRIOR APPLICATION NUMBER: PCT/US98/01373  
; PRIOR FILING DATE: 1998-01-23  
; PRIOR APPLICATION NUMBER: US 60/036,713  
; PRIOR FILING DATE: 1997-01-23  
; PRIOR APPLICATION NUMBER: US 60/037,432  
; PRIOR FILING DATE: 1997-02-07  
; NUMBER OF SEQ ID NOS: 274  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 243  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-103-395-243

Query Match 7.9%; Score 15; DB 13; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 34 VYLLPRRGPRLGVR 48
Db 1 VYLLPRRGPRLGVR 15

RESULT 2
US-10-268-561-4
; Sequence 4, Application US/10268561
; Publication No. US2003014833A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: HCV Anti-Core Monoclonal Antibodies
; FILE REFERENCE: CDS0286
; CURRENT APPLICATION NUMBER: US/10/268,561
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/337453
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-268-561-4

Query Match 7.9%; Score 15; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 QIVGGVYLLPRRGPR 43
Db 1 QIVGGVYLLPRRGPR 15

RESULT 3
US-10-268-561-5
; Sequence 5, Application US/10268561
; Publication No. US2003014833A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: HCV Anti-Core Monoclonal Antibodies
; FILE REFERENCE: CDS0286
; CURRENT APPLICATION NUMBER: US/10/268,561
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/337453
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-268-561-5

Query Match 7.9%; Score 15; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 RRGPRLGVRATRKTS 53
Db 1 RRGPRLGVRATRKTS 15

RESULT 4
US-10-268-561-6
; Sequence 6, Application US/10268561
; Publication No. US2003014833A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: HCV Anti-Core Monoclonal Antibodies
; FILE REFERENCE: CDS0286
; CURRENT APPLICATION NUMBER: US/10/268,561
; CURRENT FILING DATE: 2002-10-10
; QUERY Match 7.9%; Score 15; DB 14; Length 15;
; Best Local Similarity 100.0%; Pred. No. 2.8e-06;
; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 48 ATRKTSERSQPRGR 62
Db 1 ATRKTSERSQPRGR 15

RESULT 5
US-10-268-561-9
; Sequence 9, Application US/10268561
; Publication No. US2003014833A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: HCV Anti-Core Monoclonal Antibodies
; FILE REFERENCE: CDS0286
; CURRENT APPLICATION NUMBER: US/10/268,561
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/337453
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-268-561-9

Query Match 7.9%; Score 15; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 77 AQPGVPWPLYGNEGC 91
Db 1 AQPGVPWPLYGNEGC 15

RESULT 6
US-10-268-561-10
; Sequence 10, Application US/10268561
; Publication No. US2003014833A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: HCV Anti-Core Monoclonal Antibodies
; FILE REFERENCE: CDS0286
; CURRENT APPLICATION NUMBER: US/10/268,561
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/337453
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-268-561-10

Query Match 7.9%; Score 15; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 86 YGNEGCGWAGWLLSP 100
```

```
Db      1 YGNEGCGWAGWLLSP 15
|||||
RESULT 7
US-10-268-561-18
; Sequence 18, Application US/10268561
; Publication No. US2003014833A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: HCV Anti-Core Monoclonal Antibodies
; FILE REFERENCE: CDS0286
; CURRENT APPLICATION NUMBER: US/10/268,561
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/337453
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-268-561-18
Query Match      7.9%; Score 15; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      166 TGNLPGCSFSIFLLA 180
|||||
Db      1 TGNLPGCSFSIFLLA 15
|||||
RESULT 8
US-10-268-569-4
; Sequence 4, Application US/10268569
; Publication No. US20030152965A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: HCV Core Protein Sequences
; FILE REFERENCE: CDS-0288
; CURRENT APPLICATION NUMBER: US/10/268,569
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 60/347,303
; PRIOR FILING DATE: 2001-11-11
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-268-569-4
Query Match      7.9%; Score 15; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      29 QIVGGVYLLPRGPR 43
|||||
Db      1 QIVGGVYLLPRGPR 15
|||||
RESULT 9
US-10-268-569-5
; Sequence 5, Application US/10268569
; Publication No. US20030152965A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: HCV Core Protein Sequences
; FILE REFERENCE: CDS-0288
; CURRENT APPLICATION NUMBER: US/10/268,569
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 60/347,303
; PRIOR FILING DATE: 2001-11-11
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-268-569-5
Query Match      7.9%; Score 15; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      48 ATRKTSERSQPRGR 62
|||||
Db      1 ATRKTSERSQPRGR 15
|||||
RESULT 10
US-10-268-569-6
; Sequence 6, Application US/10268569
; Publication No. US20030152965A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: HCV Core Protein Sequences
; FILE REFERENCE: CDS-0288
; CURRENT APPLICATION NUMBER: US/10/268,569
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 60/347,303
; PRIOR FILING DATE: 2001-11-11
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-268-569-6
Query Match      7.9%; Score 15; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      39 RRGPRLGVRATRKTS 53
|||||
Db      1 RRGPRLGVRATRKTS 15
|||||
RESULT 11
US-10-268-569-9
; Sequence 9, Application US/10268569
; Publication No. US20030152965A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: HCV Core Protein Sequences
; FILE REFERENCE: CDS-0288
; CURRENT APPLICATION NUMBER: US/10/268,569
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 60/347,303
; PRIOR FILING DATE: 2001-11-11
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-268-569-9
Query Match      7.9%; Score 15; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      77 AOPGYWPPLYCNEGC 91
|||||
```

```
Db      1 AQPQYPMPLVGNQEC 15

RESULT 12
US-10-268-569-10
; Sequence 10, Application US/10268569
; Publication No. US20030152965A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: HCV Core Protein Sequences
; FILE REFERENCE: CDS-0288
; CURRENT APPLICATION NUMBER: US/10/268,569
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 60/347,303
; PRIOR FILING DATE: 2001-11-11
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-268-569-10

Query Match      7.9%; Score 15; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      86 YGNECGGWAGWLLSP 100
Db      1 YGNECGGWAGWLLSP 15

RESULT 13
US-10-268-569-18
; Sequence 18, Application US/10268569
; Publication No. US20030152965A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: HCV Core Protein Sequences
; FILE REFERENCE: CDS-0288
; CURRENT APPLICATION NUMBER: US/10/268,569
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 60/347,303
; PRIOR FILING DATE: 2001-11-11
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-268-569-18

Query Match      7.9%; Score 15; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      166 TGNLPGCSFSIFLLA 180
Db      1 TGNLPGCSFSIFLLA 15

RESULT 14
US-10-651-165-1
; Sequence 1, Application US/10651165
; Publication No. US20040047877A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
```

```
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 15
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-651-165-1

Query Match      7.9%; Score 15; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      5 PKPQKTKRNTNRRP 19
Db      1 PKPQKTKRNTNRRP 15

RESULT 15
US-10-621-675-134
; Sequence 134, Application US/10621675
; Publication No. US20050049398A1
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-11
; CURRENT APPLICATION NUMBER: US/10/621,675
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US/09/576,824A
; PRIOR FILING DATE: 08/723,425
; PRIOR APPLICATION NUMBER: 1996-09-30
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 134
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-621-675-134

Query Match      7.9%; Score 15; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      5 PKPQKTKRNTNRRP 19
Db      1 PKPQKTKRNTNRRP 15

Search completed: July 15, 2005, 10:40:08
Job time : 162 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2005, 10:07:50 ; Search time 40 Seconds  
(without alignments)  
459.435 Million cell updates/sec

Title: US-09-084-691B-206  
Perfect score: 191  
Sequence: 1 MSTLPKPKRKTNRNRRPT.....CSFSIFLLALLSCLTPASA 191

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0  
Maximum DB seq length: 15

Post-processing: Listing first 45 summaries

Database : PIR 79: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5	2.6	13	2 A38929	glutathione peroxi
2	5	2.6	14	2 S12171	H+-transporting tw
3	5	2.6	15	2 PNO148	omega-gliadin 3
4	5	2.6	15	2 B35389	urease (EC 3.5.1.5
5	4	2.1	4	2 A32480	achatin-I - giant
6	4	2.1	5	2 B61168	cocoonase (EC 3.4.
7	4	2.1	7	2 E61491	seed protein ws-5
8	4	2.1	7	2 B33541	hypothetical prote
9	4	2.1	9	2 G58502	kidney and bladder
10	4	2.1	9	2 PC7076	spectrin alpha cha
11	4	2.1	10	1 ECLQ4M	tachykinin IV - mi
12	4	2.1	10	2 S24190	tryptase (EC 3.4.2
13	4	2.1	10	2 T17075	cytochrome-c oxida
14	4	2.1	10	2 T14212	cytochrome-c oxida
15	4	2.1	11	1 A60654	substance P - guin
16	4	2.1	11	1 SPHO	substance P - hors
17	4	2.1	11	2 E41476	probable antigen 5
18	4	2.1	11	2 PNO044	protein kinase C i
19	4	2.1	11	2 D41946	T-cell receptor ga
20	4	2.1	11	2 I52304	gene rSSTR4 protei
21	4	2.1	11	2 A38590	transforming prote
22	4	2.1	11	2 A34662	Achatina cardio-ex
23	4	2.1	11	2 T12244	cytochrome-c oxida
24	4	2.1	11	2 T12248	cytochrome-c oxida
25	4	2.1	11	2 T17081	cytochrome-c oxida
26	4	2.1	11	2 T17078	cytochrome-c oxida
27	4	2.1	11	4 PC2124	aminotransferase c
28	4	2.1	12	2 S70337	napin small chain
29	4	2.1	13	2 S01043	glutamate-ammonia

30	4	2.1	14	1 QMVRMM	mastoparan M - hor
31	4	2.1	14	1 QMWAVV	mastoparan - yello
32	4	2.1	14	2 S74128	superoxide dismuta
33	4	2.1	14	2 PA0094	hyocyanine (6s)-di
34	4	2.1	14	2 S58426	spermadhesin AWN h
35	4	2.1	14	2 S66234	sperm motility inh
36	4	2.1	14	2 B60683	malate dehydrogena
37	4	2.1	14	2 D45474	thrombospondin 2 -
38	4	2.1	15	2 S24159	leukocyte elastase
39	4	2.1	15	2 I52734	gene c-Ki-ras prot
40	4	2.1	15	2 S47367	T-cell antigen rec
41	4	2.1	15	2 I65478	c-Ki-ras - hamster
42	4	2.1	15	2 PNO052	pyruvate kinase (E
43	4	2.1	15	2 S59492	formate dehydrogen
44	4	2.1	15	2 PNO164	hyocyanine (6S)-d
45	4	2.1	15	2 A49177	22K protein pl, mi

ALIGNMENTS

RESULT 1

A38929  
Glutathione peroxidase (EC 1.11.1.9) - sheep (fragment)  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C:Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 09-Jul-2004  
C:Accession: A38929  
R:Gething, P.; Dyal, D.; Crews, B.  
Arch. Biochem. Biophys. 294, 511-518, 1992  
A:Title: Selenium-dependent glutathione peroxidases from ovine and bovine erythrocytes  
A:Reference number: S21712; MUID:92231574; PMID:1567207  
A:Accession: A38929  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-13 <GET>  
A:Cross-references: UNIPROT:Q7M355  
C:Superfamily: glutathione peroxidase  
C:Keywords: oxidoreductase

Query Match 2.6%; Score 5; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 4.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 148 AAALA 152  
Db 2 AAALA 6

RESULT 2

S12171  
H+-transporting two-sector ATPase (EC 3.6.3.14) lipid-binding protein - fission yeast (S  
C:Species: mitochondrion Schizosaccharomyces pombe  
C:Date: 18-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S12171  
R:Massardo, D.R.  
Nucleic Acids Res. 18, 6429, 1990  
A:Title: Nucleotide sequence of the genes encoding tRNA (his), tRNA (pro) and tRNA (gin) in  
A:Reference number: S12171; MUID:91057135; PMID:2243789  
A:Accession: S12171  
A:Status: preliminary; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-14 <MAS>  
A:Cross-references: UNIPROT:P21535; EMBL:X54552; NID:G13659; PIDN:CAA38422.1; PID:G11923

Query Match 2.6%; Score 5; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 126 LTCGF 130

```

Db          2 LTCGF 6
|||||
RESULT 3
PN0148
omega-gliadine 3 - Aegilops longissima (fragment)
C:Species: Aegilops longissima
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PN0148
R:Odintsova, T.I.; Egorov, T.A.
Biokhimiia 55, 509-516, 1990
A:Title: N-terminal sequences of omega-gliadins of Aegilops longissima: On the origin of
A:Reference number: PN0146; MUID:90283493; PMID:2354218
A:Accession: PN0148
A:Molecule type: protein
A:Residues: 1-15 <ODI>
A:Experimental source: strain K-202
A:Note: 11-Gln was also found

Query Match          2.6%; Score 5; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 LSPRG 102
|||||
Db 4 LSPRG 8

RESULT 4
B35389
urease (EC 3.5.1.5) 15K chain - Morganella morganii (fragment)
C:Species: Morganella morganii
C:Date: 31-Aug-1990 #sequence_revision 31-Aug-1990 #text_change 09-Jul-2004
C:Accession: B35389
R:Hu, L.T.; Nicholson, E.B.; Jones, B.D.; Lynch, M.J.; Mobley, H.L.T.
J. Bacteriol. 172, 3073-3080, 1990
A:Title: Morganella morganii urease: purification, characterization, and isolation of ge
A:Reference number: A35389; MUID:90264298; PMID:2345135
A:Accession: B35389
A:Molecule type: protein
A:Status: preliminary
A:Residues: 1-15 <HUA>
A:Cross-references: UNIPROT:P17338
C:Keywords: hydrolase

Query Match          2.6%; Score 5; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 PLGGV 147
|||||
Db 8 PLGGV 12

RESULT 5
A32480
achatin-I - giant African snail
N:Contains: achatin-II
C:Species: Achatina fulica (giant African snail)
C:Date: 12-Oct-1989 #sequence_revision 12-Oct-1989 #text_change 09-Jul-2004
C:Accession: A32480
R:Kamatani, Y.; Minakata, H.; Kenny, P.T.M.; Iwashita, T.; Watanabe, K.; Funase, K.; Sur
Biochem. Biophys. Res. Commun. 160, 1015-1020, 1989
A:Title: Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina fulica feru
A:Reference number: A32480; MUID:89273551; PMID:2597281
A:Accession: A32480
A:Molecule type: protein
A:Residues: 1-4 <KAW>
A:Cross-references: UNIPROT:P35904
A:Note: Stereochemistry of the active form confirmed by chemical synthesis
R:Ishida, T.; In, Y.; Inoue, M.; Yasuda-Kamatani, Y.; Minakata, H.; Iwashita, T.; Nomoto
FEBS Lett. 307, 253-256, 1992

A:Title: Effect of the D-Phe(2) residue on molecular conformation of an endogenous neuro
(H-Gly-Phe-Ala-Asp-OH).
A:Reference number: A44691; MUID:92354723; PMID:1644179
A:Contents: annotation; X-ray crystallography, 0.85 angstroms
A:Note: achatin-II has L-phenylalanine
C:Keywords: D-amino acid
F:2/Modified site: D-phenylalanine (Phe) #status experimental

Query Match          2.1%; Score 4; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 GFAD 132
|||||
Db 1 GFAD 4

RESULT 6
B61168
cocoanase (EC 3.4.21.-) - Chinese oak silkworm (fragment)
C:Species: Antherea pernyi (Chinese oak silkworm)
C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 07-May-1999
C:Accession: B61168
R:Kramer, K.J.; Felsted, R.L.; Law, J.H.
J. Biol. Chem. 248, 3021-3028, 1973
A:Title: Cocoanase. V. Structural studies on an insect serine protease.
A:Reference number: A61168; MUID:73166540; PMID:4735570
A:Accession: B61168
A:Molecule type: protein
A:Residues: 1-5 <KRA>
C:Keywords: hydrolase; serine proteinase; zymogen
F:1-5/Product: cocoanase (fragment) #status experimental <MAT>

Query Match          2.1%; Score 4; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 IVGG 33
|||||
Db 1 IVGG 4

RESULT 7
E61491
seed protein ws-5 - winged bean (fragment)
C:Species: Psophocarpus tetragonolobus (winged bean)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 07-Oct-1994
C:Accession: E61491
R:Hirano, H.
J. Protein Chem. 8, 115-130, 1989
A:Title: Microsequence analysis of winged bean seed proteins electroblooded from two-dim
A:Reference number: A61491; MUID:89351606; PMID:2765119
A:Accession: E61491
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-7 <HIR>
C:Keywords: glycoprotein; seed

Query Match          2.1%; Score 4; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 PGGG 28
|||||
Db 4 PGGG 7

RESULT 8
B33541
hypothetical protein (T1 5' region) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 21-Feb-1990 #sequence_revision 21-Feb-1990 #text_change 30-Sep-1993
C:Accession: B33541

```

R;Klenenz, R.; Hoffmann, S.; Werenskiold, A.K.  
Proc. Natl. Acad. Sci. U.S.A. 86, 5708-5712, 1989  
A;Title: Serum- and oncoprotein-mediated induction of a gene with sequence similarity to  
A;Reference number: A33541; MUID:89345536; PMID:2527364  
A;Accession: B33541  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-7 <KLE>  
A;Cross-references: GB:M24843

Query Match 2.1%; Score 4; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 RRRS 116  
|||  
DB 2 RRRS 5

RESULT 9  
G58502  
kidney and bladder stone protein - unidentified bacterium (fragment)  
C;Species: unidentified bacterium  
C;Date: 07-Feb-1997 #sequence\_revision 07-Feb-1997 #text\_change 09-Jul-2004  
C;Accession: G58502  
R;Binette, J.P.; Binette, M.B.  
submitted to the Protein Sequence Database, October 1996  
A;Description: The proteins of kidney and gallbladder stones.  
A;Reference number: A58501  
A;Accession: G58502  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-9 <BIN>  
A;Cross-references: UNIPROT:Q7M139  
A;Experimental source: human kidney stone, bladder stone  
A;Note: a secondary sequence AAKENPXD was also found

Query Match 2.1%; Score 4; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DVKF 24  
|||  
DB 4 DVKF 7

RESULT 10  
PC7076  
spectrin alpha chain, non-erythroid - mouse (fragment)  
N;Alternate names: fodrin alpha chain  
C;Species: Mus musculus (house mouse)  
C;Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: PC7076  
R;Tsugita, A.; Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Matsui, T.; Watanabe, Y.;  
Electrophoresis 21, 1853-1871, 2000  
A;Title: Proteome analysis of mouse brain: Two-dimensional electrophoresis profiles of  
A;Reference number: PC7072  
A;Accession: PC7076  
A;Molecule type: protein  
A;Residues: 1-9 <TSU>  
A;Cross-references: UNIPROT:P16546  
A;Experimental source: strain C57BL/6Cr Slc, male; brain, striatum  
C;Keywords: brain

Query Match 2.1%; Score 4; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 ATRK 51  
|||  
DB 3 ATRK 6

RESULT 11  
BCLQ4M  
tachykinin IV - migratory locust  
N;Alternate names: locusttachykinin IV  
C;Species: Locusta migratoria (migratory locust)  
C;Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 09-Jul-2004  
C;Accession: B60073  
R;Schoofs, L.; Holman, G.M.; Hayes, T.K.; Kochansky, J.P.; Nachman, R.J.; De Loof, A.  
Regul. Pept. 31, 199-212, 1990  
A;Title: Locusttachykinin III and IV: two additional insect neuropeptides with homology  
A;Reference number: A60073; MUID:91219696; PMID:2132575  
A;Accession: B60073  
A;Molecule type: protein  
A;Residues: 1-10 <SCH>  
A;Cross-references: UNIPROT:P30250  
C;Superfamily: tachykinin  
C;Keywords: amidated carboxyl end; neuropeptide; tachykinin  
F;10/Modified site: amidated carboxyl end (Arg) #status experimental

Query Match 2.1%; Score 4; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 153 HGVR 156  
|||  
DB 7 HGVR 10

RESULT 12  
S24190  
tryptase (EC 3.4.21.59) - bovine (fragment)  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 19-Mar-1997 #sequence\_revision 18-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: S24190  
R;Fiorucci, L.; Erba, F.; Ascoli, F.  
Biochem. Hoppe-Seyler 373, 483-490, 1992  
A;Title: Bovine tryptase: purification and characterization.  
A;Reference number: S24190; MUID:92384956; PMID:1515079  
A;Accession: S24190  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-10 <FIO>  
A;Cross-references: UNIPROT:Q8HYJ2  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; serine proteinase; zymogen

Query Match 2.1%; Score 4; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 IVGG 33  
|||  
DB 1 IVGG 4

RESULT 13  
T17075  
cytochrome-c oxidase (EC 1.9.3.1) chain I - Chamaeleo fischeri mitochondrion (fragment)  
C;Species: mitochondrion Chamaeleo fischeri  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T17075  
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Papenfuss, T.J.  
J. Mol. Evol. 44, 660-674, 1997  
A;Title: Evolutionary shifts in three major structural features of the mitochondrial gen  
A;Reference number: Z18674; MUID:97315309; PMID:9169559  
A;Accession: T17075  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-10 <MAC>  
A;Cross-references: UNIPROT:O79912; EMBL:U82688; NID:G3603112; PID:G3603115; PIDN:AAC622

C;Keywords: mitochondrion; oxidoreductase

Query Match 2.1%; Score 4; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 WLLS 99  
|||  
Db 6 WLLS 9

#### RESULT 14

T14212  
cytochrome-c oxidase (EC 1.9.3.1) chain I - Uromastix acanthinurus mitochondrion (fragment)  
C;Species: mitochondrion Uromastix acanthinurus  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: T14212  
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.  
Mol. Biol. Evol. 14, 91-104, 1997  
A;Title: Two novel gene orders and the role of light-strand replication in rearrangement  
A;Reference number: Z17789; MUID:97153826; PMID:9000757  
A;Accession: T14212  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-10 <MAC>  
A;Cross-references: UNIPROT:P92762; EMBL:U71325; NID:g1753264; PID:g1753267; PIDN:AAC622  
C;Genetics:  
A;Genome: mitochondrion  
A;Note: COI  
C;Keywords: mitochondrion; oxidoreductase

Query Match 2.1%; Score 4; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 WLLS 99  
|||  
Db 6 WLLS 9

#### RESULT 15

A60654  
substance P - guinea pig  
C;Species: Cavia porcellus (guinea pig)  
C;Date: 14-May-1993 #sequence\_revision 27-Jun-1994 #text\_change 09-Jul-2004  
C;Accession: A60654  
R;Murphy, R.  
Neuropeptides 14, 105-110, 1989  
A;Title: Primary amino acid sequence of guinea-pig substance P.  
A;Reference number: A60654; MUID:90044685; PMID:2478925  
A;Accession: A60654.  
A;Molecule type: protein  
A;Residues: 1-11 <MUR>  
A;Cross-references: UNIPROT:P01290  
C;Superfamily: substance P precursor  
C;Keywords: amidated carboxyl end; neuropeptide; tachykinin  
F;11/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 2.1%; Score 4; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 3.9e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PKPQ 8  
|||  
Db 2 PKPQ 5

Search completed: July 15, 2005, 10:27:21  
Job time : 42 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2005, 09:59:29 ; Search time 173 Seconds  
(without alignments)  
565.359 Million cell updates/sec

Title: US-09-084-691B-206  
Perfect score: 191  
Sequence: 1 MSTLPKPKQKTKRNTNRPT.....CSFSIFLLALLSCLTTPASA 191

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0  
Maximum DB seq length: 15

Post-processing: Listing first 45 summaries

Database : Uniprot\_03: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	13	6.8	13	2	Q81766 hepatitis c
2	13	6.8	13	2	Q81771 hepatitis c
3	13	6.8	13	2	Q81790 hepatitis c
4	13	6.8	13	2	Q81792 hepatitis c
5	13	6.8	13	2	Q81796 hepatitis c
6	9	4.7	13	2	Q81761 hepatitis c
7	9	4.7	13	2	Q81762 hepatitis c
8	9	4.7	13	2	Q81763 hepatitis c
9	9	4.7	13	2	Q81764 hepatitis c
10	9	4.7	13	2	Q81765 hepatitis c
11	9	4.7	13	2	Q81767 hepatitis c
12	9	4.7	13	2	Q81768 hepatitis c
13	9	4.7	13	2	Q81769 hepatitis c
14	9	4.7	13	2	Q81772 hepatitis c
15	9	4.7	13	2	Q81773 hepatitis c
16	9	4.7	13	2	Q81774 hepatitis c
17	9	4.7	13	2	Q81775 hepatitis c
18	9	4.7	13	2	Q81777 hepatitis c
19	9	4.7	13	2	Q81778 hepatitis c
20	9	4.7	13	2	Q81779 hepatitis c
21	9	4.7	13	2	Q81780 hepatitis c
22	9	4.7	13	2	Q81782 hepatitis c
23	9	4.7	13	2	Q81783 hepatitis c
24	9	4.7	13	2	Q81784 hepatitis c
25	9	4.7	13	2	Q81785 hepatitis c
26	9	4.7	13	2	Q81786 hepatitis c
27	9	4.7	13	2	Q81787 hepatitis c
28	9	4.7	13	2	Q81788 hepatitis c
29	9	4.7	13	2	Q81789 hepatitis c
30	9	4.7	13	2	Q81791 hepatitis c
31	9	4.7	13	2	Q81793 hepatitis c

32 9 4.7 13 2 Q81794 hepatitis c  
33 9 4.7 13 2 Q81795 hepatitis c  
34 9 4.7 13 2 Q81797 hepatitis c  
35 9 4.7 13 2 Q81798 hepatitis c  
36 9 4.7 13 2 Q81799 hepatitis c  
37 9 4.7 13 2 Q81800 hepatitis c  
38 7 3.7 13 2 Q81781 hepatitis c  
39 5 2.6 12 2 Q8Y3L5  
40 5 2.6 12 2 Q8Y3L6  
41 5 2.6 12 2 Q8Y3M0  
42 5 2.6 12 2 Q8Y3M1  
43 5 2.6 13 1 SCX2\_TITDI  
44 5 2.6 13 2 Q8UP66  
45 5 2.6 13 2 Q7M355

## ALIGNMENTS

RESULT 1  
Q81766 .  
ID Q81766 PRELIMINARY; PRT; 13 AA.  
AC Q81766;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Polyprotein (Fragment).  
GN Name=polyprotein;  
OS Hepatitis C virus;  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92279243; PubMed=1317578;  
RT Bukh J., Purcell R.H., Miller R.H.;  
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).  
DR EMBL; M84827; AAA45684.1; -;  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR002522; HCV\_capsid; 1.  
DR Pfam; PF01543; HCV\_capsid; 1.  
KW Polyprotein.  
FT NON\_TER  
SQ SEQUENCE 13 AA; 1571 MW; 464DIDEIA42FC763 CRC64;  
Query Match 6.8%; Score 13; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 3.4e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSTLPKPKQKTKR 13  
|||  
Db 1 MSTLPKPKQKTKR 13  
RESULT 2  
Q81771 .  
ID Q81771 PRELIMINARY; PRT; 13 AA.  
AC Q81771;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Polyprotein (Fragment).  
GN Name=polyprotein;  
OS Hepatitis C virus;  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92279243; PubMed=1317578;  
RT Bukh J., Purcell R.H., Miller R.H.;

```
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR EMBL; M84834; AAA45689.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER 1
SQ SEQUENCE 13 AA; 1571 MW; 464D1DE1A42FC763 CRC64;

Query Match 6.8%; Score 13; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTLPKPQKTKR 13
   |||||
DB 1 MSTLPKPQKTKR 13
   |||||

RESULT 3
Q81790 ID Q81790 PRELIMINARY; PRT; 13 AA.
AC Q81790;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polyprotein (Fragment).
GN Name=polyprotein;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92279243; PubMed=1317578;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR EMBL; M84864; AAA45707.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER 1
SQ SEQUENCE 13 AA; 1571 MW; 464D1DE1A42FC763 CRC64;

Query Match 6.8%; Score 13; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTLPKPQKTKR 13
   |||||
DB 1 MSTLPKPQKTKR 13
   |||||

RESULT 4
Q81792 ID Q81792 PRELIMINARY; PRT; 13 AA.
AC Q81792;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polyprotein (Fragment).
GN Name=polyprotein;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92279243; PubMed=1317578;
```

```
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR EMBL; M84835; AAA45709.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER 1
SQ SEQUENCE 13 AA; 1571 MW; 464D1DE1A42FC763 CRC64;

Query Match 6.8%; Score 13; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTLPKPQKTKR 13
   |||||
DB 1 MSTLPKPQKTKR 13
   |||||

RESULT 5
Q81796 ID Q81796 PRELIMINARY; PRT; 13 AA.
AC Q81796;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polyprotein (Fragment).
GN Name=polyprotein;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92279243; PubMed=1317578;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR EMBL; M84837; AAA45713.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER 1
SQ SEQUENCE 13 AA; 1571 MW; 464D1DE1A42FC763 CRC64;

Query Match 6.8%; Score 13; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTLPKPQKTKR 13
   |||||
DB 1 MSTLPKPQKTKR 13
   |||||

RESULT 6
Q81761 ID Q81761 PRELIMINARY; PRT; 13 AA.
AC Q81761;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polyprotein (Fragment).
GN Name=polyprotein;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RX MEDLINE=92279243; PubMed=1317578;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR EMBL; M84842; AAA45679.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER
SQ SEQUENCE 13 AA; 1572 MW; 464F97E1A42FC763 CRC64;

Query Match 4.7%; Score 9; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PKPQKTKR 13
DB 5 PKPQKTKR 13

RESULT 7
Q81762 ID Q81762 PRELIMINARY; PRT; 13 AA.
AC Q81762;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polyprotein (Fragment)
GN Name=polyprotein;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92279243; PubMed=1317578;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR EMBL; M84841; AAA45680.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER
SQ SEQUENCE 13 AA; 1572 MW; 464F97E1A42FC763 CRC64;

Query Match 4.7%; Score 9; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PKPQKTKR 13
DB 5 PKPQKTKR 13

RESULT 8
Q81763 ID Q81763 PRELIMINARY; PRT; 13 AA.
AC Q81763;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polyprotein (Fragment)
GN Name=polyprotein;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92279243; PubMed=1317578;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR EMBL; M84841; AAA45680.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER
SQ SEQUENCE 13 AA; 1572 MW; 464F97E1A42FC763 CRC64;

Query Match 4.7%; Score 9; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PKPQKTKR 13
DB 5 PKPQKTKR 13

RESULT 9
Q81764 ID Q81764 PRELIMINARY; PRT; 13 AA.
AC Q81764;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polyprotein (Fragment)
GN Name=polyprotein;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92279243; PubMed=1317578;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR EMBL; M84851; AAA45682.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER
SQ SEQUENCE 13 AA; 1572 MW; 464F97E1A42FC763 CRC64;

Query Match 4.7%; Score 9; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PKPQKTKR 13
DB 5 PKPQKTKR 13

RESULT 10
Q81765 ID Q81765 PRELIMINARY; PRT; 13 AA.
AC Q81765;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polyprotein (Fragment)
GN Name=polyprotein;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=92279243; PubMed=1317578;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR EMBL; M84839; AAA45681.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER
SQ SEQUENCE 13 AA; 1572 MW; 464F97E1A42FC763 CRC64;

Query Match 4.7%; Score 9; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PKPQKTKR 13
DB 5 PKPQKTKR 13

RESULT 9
Q81764 ID Q81764 PRELIMINARY; PRT; 13 AA.
AC Q81764;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polyprotein (Fragment)
GN Name=polyprotein;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92279243; PubMed=1317578;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR EMBL; M84851; AAA45682.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER
SQ SEQUENCE 13 AA; 1572 MW; 464F97E1A42FC763 CRC64;

Query Match 4.7%; Score 9; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PKPQKTKR 13
DB 5 PKPQKTKR 13

RESULT 10
Q81765 ID Q81765 PRELIMINARY; PRT; 13 AA.
AC Q81765;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polyprotein (Fragment)
GN Name=polyprotein;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
```

```

RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92279243; PubMed=1317578;
RA  Bukh J., Purcell R.H., Miller R.H.;
RT  "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL  Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR  EMBL; M84859; AAA45683.1; -
DR  GO; GO:0019028; C:viral capsid; IEA.
DR  GO; GO:0005198; F:structural molecule activity; IEA.
DR  InterPro; IPR002522; HCV capsid.
DR  Pfam; PF01543; HCV_capsid; 1.
KW  Polyprotein.
FT  NON_TER
SQ  SEQUENCE 13 AA; 1572 MW; 464F97E1A42FC763 CRC64;

Query Match      4.7%; Score 9; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  5 PKPQKTKR 13
Db   |||||
     5 PKPQKTKR 13

RESULT 11
Q81767
ID  Q81767 PRELIMINARY; PRT; 13 AA.
AC  Q81767;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  Polyprotein (Fragment).
GN  Name=polyprotein;
OS  Hepatitis C virus.
OC  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC  Hepacivirus.
OX  NCBI_TaxID=11103;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92279243; PubMed=1317578;
RA  Bukh J., Purcell R.H., Miller R.H.;
RT  "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL  Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR  EMBL; M84849; AAA45685.1; -
DR  GO; GO:0019028; C:viral capsid; IEA.
DR  GO; GO:0005198; F:structural molecule activity; IEA.
DR  InterPro; IPR002522; HCV capsid.
DR  Pfam; PF01543; HCV_capsid; 1.
KW  Polyprotein.
FT  NON_TER
SQ  SEQUENCE 13 AA; 1572 MW; 464F97E1A42FC763 CRC64;

Query Match      4.7%; Score 9; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  5 PKPQKTKR 13
Db   |||||
     5 PKPQKTKR 13

RESULT 12
Q81768
ID  Q81768 PRELIMINARY; PRT; 13 AA.
AC  Q81768;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  Polyprotein (Fragment).
GN  Name=polyprotein;
OS  Hepatitis C virus.
OC  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC  Hepacivirus.

```

```

OX  NCBI_TaxID=11103;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92279243; PubMed=1317578;
RA  Bukh J., Purcell R.H., Miller R.H.;
RT  "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL  Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR  EMBL; M84855; AAA45686.1; -
DR  GO; GO:0019028; C:viral capsid; IEA.
DR  GO; GO:0005198; F:structural molecule activity; IEA.
DR  InterPro; IPR002522; HCV capsid.
DR  Pfam; PF01543; HCV_capsid; 1.
KW  Polyprotein.
FT  NON_TER
SQ  SEQUENCE 13 AA; 1572 MW; 464F97E1A42FC763 CRC64;

Query Match      4.7%; Score 9; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  5 PKPQKTKR 13
Db   |||||
     5 PKPQKTKR 13

RESULT 13
Q81769
ID  Q81769 PRELIMINARY; PRT; 13 AA.
AC  Q81769;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  Polyprotein (Fragment).
GN  Name=polyprotein;
OS  Hepatitis C virus.
OC  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC  Hepacivirus.
OX  NCBI_TaxID=11103;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92279243; PubMed=1317578;
RA  Bukh J., Purcell R.H., Miller R.H.;
RT  "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL  Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR  EMBL; M84856; AAA45687.1; -
DR  GO; GO:0019028; C:viral capsid; IEA.
DR  GO; GO:0005198; F:structural molecule activity; IEA.
DR  InterPro; IPR002522; HCV capsid.
DR  Pfam; PF01543; HCV_capsid; 1.
KW  Polyprotein.
FT  NON_TER
SQ  SEQUENCE 13 AA; 1559 MW; 465365E1A42FC763 CRC64;

Query Match      4.7%; Score 9; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  5 PKPQKTKR 13
Db   |||||
     5 PKPQKTKR 13

RESULT 14
Q81772
ID  Q81772 PRELIMINARY; PRT; 13 AA.
AC  Q81772;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  Polyprotein (Fragment).
GN  Name=polyprotein;
OS  Hepatitis C virus.
OC  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

```

```

OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92279243; PubMed=1317578;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR EMBL; M84833; AAA45690.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER
SQ SEQUENCE 13 AA; 1572 MW; 464F97E1A42FC763 CRC64;

Query Match 4.7%; Score 9; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PKPQKTKR 13
DB 5 PKPQKTKR 13

RESULT 15
Q81773
ID Q81773 PRELIMINARY; PRT; 13 AA.
AC Q81773;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polyprotein (Fragment).
GN Name=polyprotein;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92279243; PubMed=1317578;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
DR EMBL; M84833; AAA45691.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER
SQ SEQUENCE 13 AA; 1572 MW; 464F97E1A42FC763 CRC64;

Query Match 4.7%; Score 9; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PKPQKTKR 13
DB 5 PKPQKTKR 13

```

Search completed: July 15, 2005, 10:26:34  
Job time : 174 secs

***This Page Blank (uspto)***